



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 162981

TO: **Phuong Bui**
Art Unit: **1638**
Location: **REM-2A15&2C18**
Serial Number: **10/660226**

Wednesday, August 31, 2005

From: **Beverly Shears**
Location: **Biotech-Chem Library**
REM 1A54
Phone: **571-272-2528**
beverly.shears@uspto.gov

Search Notes

Protein Sequence Searches – February 2005

All of the sequence databases on ABSS have recently been updated.

- Please note that the curators of the UniProt database have purged some temporary accession numbers from the most recent version of UniProt. These sequences have been assigned new permanent accession numbers. The new UniProt record may not contain the previous temporary accession number.
- If you encounter an accession number from an older search run against UniProt (results file extension **.rup**) that can no longer be found in the database, the permanent record with the new accession number can be found by searching the old accession number in the UniProt Protein Archive database (uniPARC) at:

<http://www.pir.uniprot.org/database/archive.shtml>

If you have any questions regarding this information or your results, please contact any STIC searcher.

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 25, 2005, 12:56:51 ; Search time 5404 Seconds
(without alignments)
10759.855 Million cell up

Title: US-10-660-226-9

Perfect score:

Sequence: 1 ccgccaccagctaccctgcc.....aaaaaaaaaaaaaaaaaaaaa 1200

Scoring table: IDENTITY NUC

Scoring code: IDENTIFY_NOC
Gapop 10.0 : Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Minimum DB seq length: 9
Maximum DB seq length: 200000000

Post-processing: Minimum March 0%

Fast-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl: *

1: qb ba:*

2: gb-htq:

3: qb-in: *

4: qb_om: 4

5: qb ov: *

6: gqb_pat:

7: gb_ph: *

8: gb pl: *

9: gb_pr: 4

10: gb_ro:

11: gb_ests

12: gb_sy:

13: gb_un:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query #		DB	ID	Description
		Match	Length			
1	574	47.8	1061	8	BT009529	Triticum
2	501.2	41.8	1918	8	AK109730	Oryza sat
3	423.4	35.3	1356	8	AK066687	Oryza sat
4	364.4	30.4	1323	8	BT009514	Triticum
5	242.4	20.2	1293	8	LESHKXNP	BT012939
6	242.4	20.2	1303	8	BT012939	Lycopersi
7	237.8	19.8	1300	8	AK118899	Arabidops
8	222.6	18.6	1167	8	AK070318	Oryza sat
9	193	16.1	903	8	BT005291	Arabidops
10	193	16.1	1294	8	AK117791	Arabidops
11	185.8	15.5	1016	8	AY596190	Cucumis s
12	164.4	13.7	2127	8	AK067948	Oryza sat
13	157	13.1	309	6	AK315616	Sequence
14	120.4	10.0	111777	2	AP004089	Oryza sat
15	120.4	10.0	154796	8	AP004774	Oryza sat
16	120.4	10.0	155168	8	AP005848	Oryza sat
17	115.8	9.7	137319	8	AP003513	Oryza sat
18	85.4	7.1	137740	1	D90900	Synechocyst
19	84.8	7.1	1843	8	BT012407	Arabidops

Db 118 GAACTTGCATAACTCGTTGACGATGCCCTCTTTGTAAGAGAAATACAGAGAGTTCT 177
Qy 407 GTTCTACTTTGAACGGAGGTGTATTTACCTAGTAGGAATGATGGTTCTCTGAAAAAGTAC 466
Db 178 TTTCCAGTTTGAACGGTCGGTGCATCTACCTAGTTGGAATGATGGTTCTCGGGGAAAAAGTAC 237
Qy 467 TGTGGGGAAGATTATCTCTGAAGTCTTGGTTTATTCGTTCTTTGATAGTGACAAAGTTAGT 526
Db 238 GGTGGGGAAGATCTTGGCTGAAGTTTGGGTTATTCATTTCTCGACAGTGATAAATTGTT 297
Qy 527 GGAGCAAGCTGTGTGGAATGCCATCAGTTGCCCAATATTCAAGGTCCATAGTGAAGCCTT 586
Db 298 CGAACAAGCTGTGTGGCATGCCCTTCAGTTGCTCAAAATTTTCAAGGTTTCATAGTGAAGCCTT 357
Qy 587 CTTTCCGGATATGAGAGTAGTGTCTTGAGAGATTTCCTCATCGCAGATTAGTTGTT 646
Db 358 CTTTCAGAGATAATGAGAGTAGTGTCTTGAGGAGATTTCCTCAATCGCGCGATTAGTTGT 417
Qy 647 TGCACCCGAGGTGGTCTGTTATCCGACCAATTAACTCGAGATATATCAAGAGGGGCT 706
Db 418 TGTCTACTGGAGTGGTGTCTGTTATCCGACAGTTAACTCGAAAAATATGAAAGGGCT 477
Qy 707 ATCTGTTGGTTAGATGTGCCCTTGGATGCTCTTGTCTAGGCGGTATTGCTAAAGTGGGAC 766
Db 478 ATCTGTTGGTTAGATGTGCCCTTGGAGCTCTTGGCAAGCGGTATTGCTAAAGTGGGAC 537
Qy 767 TGCCTCTGCTCTTCTGCGACCAACCATCTGGTGATCCGTAGCGAATGGCCCTTTTCTAA 826
Db 538 TGCCTCGCTCTCTTCTAGATCAACCATCCGGTGATCATACACAATGGCCCTTTTCGAA 597
Qy 827 GCTCAGCATGCTTGACACAAAGGGGTGATGCTTATGCAAAATCGAGATGTAAGGTTTC 886
Db 598 ACTCAGCATGCTCGCGGACAAAGGGGCGATGCTTATGCAAAATCGATGTCAGATTTTC 657
Qy 887 TCTGGAAGAGATTGTCATGTAACAAGSTCATGATGATGCTCTTAAGCTGACACCTACTGA 946
Db 658 TCTCGAAGAGATCGCATCTAAGCTGGGTCTATGACGACGCTCTTAAGCTGACACCGATTGA 717
Qy 947 TATTGCAATTGATGCTACTTCAAGATCGAGAGCTTCGTATCGAGCACACTGCTGATAG 1006
Db 718 TATTGCTCTCGATGCTCTCCACAAGATCGAGAGCTTTGTGTCGAAGACACCGCT----- 772
Qy 1007 TTCAGCTAGCGCGCCCAAGCTGAGTCGCGAGATCCAGAGGATACAGACCTTGTAGAACCT 1066
Db 773 -----GTCGCCGACTCAAAACGGAATCGCAATCTCAAAGGATGCATACCTTTAGATAT 828
Qy 1067 TAATCCCTTT 1076
Db 829 GAATCCTTTT 838

RESULT 2
AKI09730
LOCUS
DEFINITION
Oryza sativa (japonica cultivar-group) cDNA clone:002-146-B12, full insert sequence.
AKI09730
ACCESSION
AKI09730.1 GI:32994939
VERSION
FLI CDNA; oligo capping.
KEYWORDS
Oryza sativa (japonica cultivar-group)
SOURCE
Oryza sativa (japonica cultivar-group)
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.
1
REFERENCE
AUTHORS
The Rice Full-length cDNA Consortium, National Institute of Agrobiological Sciences Rice Full-length cDNA Project Team:
Kikuchi S., Satoh K., Nagata T., Kawagashira N., Doi K., Kishimoto N., Yazaki J., Ishikawa M., Yamada H., Ooka H., Hotta I., Kojima K., Namiki T., Ohneda E., Yahagi W., Suzuki K., Li C., Tsunoda Y., Sugano S., Fujimura T., Shishiki T., Foundation of Advancement of International Science Genome Sequencing & Analysis Group: Ootomo Y., Tsunoda Y., Iida Y.,

Kurosaki, T., Kodama, T., Maeda, H., Kobayashi, M., Xie, Q., Lu, M., Narikawa, R., Sugiyama, A., Mizuno, K., Yokomizo S., Niikura, J., Ikeda, R., Ishibiki, J., Kawamata, M., Yoshimura, A., Miura, J., Kusumegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., RIKEN:
Hara, A., Carninci, P., Aizawa, K., Arakawa, T., Fukuda, S., Kawai, J., Hashizume, W., Hayatsu, N., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kondo, S., Konno, H., Miyazaki, A., Osato, N., Ota, Y., Saito, R., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Yoshino, M. and Hayashizaki, Y.
Collection, mapping, and annotation of over 28,000 cDNA clones from japonica rice
Science 301 (5631), 376-379 (2003)
22752273
12869764
2 (bases 1 to 1918)
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K., Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Hotta, I., Iida, J., Itoh, M., Imamura, K., Imotani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I., Kanagawa, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M., Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Kurosaki, T., Kusumegi, T., Li, C., Lu, M., Masuda, H., Matsubara, K., Matsuyama, T., Miura, J., Miyazaki, A., Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakamura, M., Namiki, T., Narikawa, R., Niikura, J., Nishi, K., Nomura, K., Numasaki, R., Ohneda, E., Ohno, M., Ohtsuki, K., Oka, M., Ooka, H., Osato, N., Ota, Y., Ootomo, Y., Ryu, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K., Shibata, K., Shinagawa, A., Shiraki, T., Shishiki, T., Sogabe, Y., Sugano, S., Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toyota, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yahagi, W., Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S. and Yoshimura, A.
Direct Submission
Submitted (27-AUG-2002) Shoshi Kikuchi, National Institute of Agrobiological Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan (E-mail:skikuchi@nias.affrc.go.jp, Tel:81-29-838-7007, Fax:81-29-838-7007)
This clone is one of the 28K full-length cDNA clones from japonica rice.
URL : http://cdna01.dna.affrc.go.jp/cDNA/
NIAS Rice Full-length cDNA Project Team: Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T. and Yamamoto, M.
PAIS Genome Sequencing & Analysis Group: Ootomo, Y., Iida, Y., Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M., Kodama, T., Kurosaki, T., Kusumegi, T., Lu, M., Maeda, H., Miura, J., Mizuno, K., Narikawa, R., Niikura, J., Oka, M., Ryu, R., Sugano, S., Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S., Yoshimura, A., Matsubara, K. and Murakami, K.
Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Iida, J., Imamura, K., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kanagawa, S., Kato, H., Kawai, J., Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Osato, N., Ota, Y., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toyota, T., Waki, K., Yasunishi, A. and Hayashizaki, Y.
Location/Qualifiers
1. :1918
/organism="Oryza sativa (japonica cultivar-group)"

/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="002-146-B12"

ORIGIN	Query Match	41.8%; Score 501.2; DB 8; Length 1918;
	Best Local Similarity	80.4%; Pred. No. 9.6e-101;
	Matches 600; Conservative 0; Mismatches 143; Indels 3; Gaps 1;	
QY	338 AGTGCATGAAACTCGCACAACTCCGTTCGACGAAGCTCTCCCTGTTGCAAGGAAATCACA	397
Db	982 AGCCACGACAGCTTGCATAACTCAGTTGATGAAGCCCTCCCTGTTAAAGAGGAAATCACA	1041
QY	398 AGAAGTCTCTGTTCTTCTACTTGAACGGGAGGTGTATTTACCTAGTAGGAAATGATGGTTCCTCG	457
Db	1042 AGAAGTCTTATCTACTTGAACGGACGGTGTATTTATTTAGTTGGAATGATGGCTCAGG	1101
QY	458 AAAAAGTACTGTGGGGAAGATTATGTCGAAGTCTTGGGTATTCGTTCTTTGATAGTGA	517
Db	1102 GAAGAGTACAGTTGCCAAGATATTAGCCGAAGTTTGGGTATTTCATTCTTTGACAGTGA	1161
QY	518 CAAGTTAGTGGACGAGCTGTTGGAAATGCCATCAGTTGCCAAATATTCAGGTCATAG	577
Db	1162 TAAGTTGGTTGAACAACGACGTTGGAATGCTCTGTGTGCCAAATATTCAGGAGCACAG	1221
QY	578 TGAAGCCTCTCTTTCGGGATAATCAGAGTAGTGTCTTTGAGAGATTGTCTCTCATGCGAGC	637
Db	1222 TGAAGCGTTTTTCAGAGATAATGAGAGTAGTGTCTTTGAGGACTGTCTCTCATGCGAGC	1281
QY	638 ATTAGTGTGTGCGACCGGAGGTGGTGTATTCGACCAATTAATCGAGATATATGAA	697
Db	1282 ATTAGTGTGTGCAACTGAGGTTGGTCCGTTATTTCGCCCAATTAATCGGAAATATATGAA	1341
QY	698 GAGGGCCATCTGTTGTTGTAGATGTCCTTGGATGCTCTTGCTAGGCGTATTGCTAA	757
Db	1342 GAAAGTCTATCTGCTGGCTGATGTCCTTTGGATGACATTGCAAGGCGTATTGCTCA	1401
QY	758 AGTGGGAATGCTCTCGTCTCTTCTGGACCAACCATCTGTGATCCGTAGCGAATGCG	817
Db	1402 AGTGGGTACTGCTCTCGTCTCTTCTGGATCAGCATCGATGATCCATACACAGCGGC	1461
QY	818 CTTTCTTAAGCTCAGCATGCTTGCA CAGCAAGGGGTGATGCTTATGCAAAATCGAGATGT	877
Db	1462 CTTCTCAAAACTCAGCATGCTTGCA CAGCAAGAGCGGATGCTTATGCAAAATGCTGATGC	1521
QY	878 AAGGTTCTCTGGRAGAGATTGCATGTAACCAAGGTCATGATGATCTCTTAAGCTGAC	937
Db	1522 TCGAGTTTCTCTTGAAGAAATTCGAACCTAAACAAGGCCATGACGATGTCTCCAAACTAAC	1581
QY	938 ACCTACTGATATTGCAATGAGTCACCTTCATAGATCGAGAGCTTCGTTCATCGAGCA---	994
Db	1582 ACCGACCGATATTGCTATTGAGGCCCTTACTCAGATTGAGATTGCTCACCAGGATTC	1641
QY	995 CACTGCTGATAGTTCAGCTAGCGACCGCAAGCTGAGTCCGAGATCCAGAGATACAGAC	1054
Db	1642 CACATCAAGTGGCCAGTTGGGGACTTAATAGTTGATTCCCAAGTCCGAAGCAAGGCAAGGC	1701
QY	1055 CTTGTAGAACCTTTAATCCCTTTGTTT	1080
Db	1702 CTTGTAGAAATTTTGGCCCTCTCTTTTTT	1727
RESULT 3		
AK066687		
LOCUS	AK066687	1356 bp mRNA linear PLN 24-JUL-2003
DEFINITION	Oryza sativa (japonica cultivar-group) cDNA clone:J013074E04, full insert sequence.	
ACCESSION	AK066687	
VERSION	AK066687.1	GI:32976705
KEYWORDS	FLI CDNA; CAP trapper.	
SOURCE	Oryza sativa (japonica cultivar-group)	
ORGANISM	Oryza sativa (japonica cultivar-group)	

REFERENCE
AUTHORS

1 The Rice Full-Length cDNA Consortium, National Institute of Agricultural Sciences Rice Full-Length cDNA Project Team: Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuka, K., Shishiki, T., Foundation of Advancement of International Science Genome Sequencing & Analysis Group: Ootomo, Y., Murakami, K., Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda, Y., Kurosaki, T., Kodama, T., Masuda, H., Kobayashi, M., Xie, Q., Lu, M., Narikawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Niikura, J., Ikeda, K., Ishibiki, J., Kawamata, M., Yoshimura, A., Miura, J., Kusumegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., RIKEN, Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashidume, W., Hayatsu, N., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kondo, S., Konno, H., Miyazaki, A., Osato, N., Ota, Y., Saito, K., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Yoshino, M., and Hayashizaki, Y.

Collection, mapping, and annotation of over 28,000 cDNA clones from japonica rice
Science 301 (5631), 376-379 (2003)

JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

2 (bases 1 to 1356)
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K., Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Hotta, I., Iida, J., Itoh, Y., Ikeda, R., Imamura, K., Imotani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I., Kikuchi, S., Kato, H., Kawagashira, N., Kawai, J., Kawamata, S., Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M., Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurahara, C., Kurosaki, T., Kusumegi, T., Li, C., Lu, M., Masuda, H., Matsubara, K., Matsuyama, T., Miura, J., Miyazaki, A., Namiki, T., Narikawa, R., Niikura, J., Nishi, K., Nomura, K., Numasaki, R., Ohneda, E., Ohno, M., Ohtsuka, K., Oka, M., Ooka, H., Ootomo, Y., Ootomo, Y., Ryu, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K., Shibata, K., Shinagawa, A., Shiraki, T., Shishiki, T., Sogabe, Y., Sugano, S., Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akashira, S., Tanaka, T., Tomaru, A., Tota, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yahagi, W., Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S. and Yoshimura, A.
Direct Submission
Submitted (05-DEC-2001) Shoshi Kikuchi, National Institute of Agricultural Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan [E-mail: skikuchi@nias.affrc.go.jp, Tel:81-29-838-7007, Fax:81-29-838-7007]
This clone is one of the 28K full-length cDNA clones from japonica rice.
URL : <http://cdna01.dna.affrc.go.jp/cDNA/>
NIAS Rice Full-length cDNA Project Team: Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuka, K., Shishiki, T. and Yamamoto, M.

TITLE
JOURNAL

Submited (05-DEC-2001) Shoshi Kikuchi, National Institute of Agricultural Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan [E-mail: skikuchi@nias.affrc.go.jp, Tel:81-29-838-7007, Fax:81-29-838-7007]
This clone is one of the 28K full-length cDNA clones from japonica rice.

COMMENT

URL : <http://cdna01.dna.affrc.go.jp/cDNA/>
NIAS Rice Full-length cDNA Project Team: Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuka, K., Shishiki, T. and Yamamoto, M.
FAIS Genome Sequencing & Analysis Group: Ootomo, Y., Iida, Y., Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M., Kodama, T., Kurosaki, T., Kusumegi, T., Lu, M., Masuda, H., Miura, J., Mizuno, K., Narikawa, R., Niikura, J., Oka, M., Ryu, R., Sugano, S., Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S., Yoshimura, A., Matsubara, K. and Murakami, K.
Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Iida, J., Imamura, K., Imotani, K., Ishii, Y.,

Itoh,M., Kagawa,I., Kanagawa,S., Katoh,H., Kawai,J.,
 Kishikawa-Hirozane,T., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
 Koya,S., Kurihara,C., MatsuYama,T., Miyazaki,A., Murata,M.,
 Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Ohno,M., Oeato,N.,
 Ota,Y., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H.,
 Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T.,
 Sogabe,Y., Tagami,M., Tagami-Takeda,Y., Tagawa,A., Takahashi,F.,
 Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T., Waki,K.,
 Yasunishi,A. and Hayashizaki,Y.
 Location/Qualifiers
 1. .1356
 /organism="Oryza sativa (japonica cultivar-group)"
 /mol_type="mRNA"
 /cultivar="Nipponbare"
 /db_xref="taxon:39947"
 /clone="J013074E04"

ORIGIN
 Query Match 35.3%; Score 423.4; DB 8; Length 1356;
 Best Local Similarity 79.6%; Pred. No. 1.8e-83;
 Matches 558; Conservative 0; Mismatches 96; Indels 47; Gaps 3;
 QY 149 GCGGGGGGCGTGGCGCTCGCGTGCAGCGCGGGCGGGCTTCGGTCCAGCGGCA 208
 Db 83 GGAGGCGGCGTGGGCGTGGCGCTGCAGTCGCGGGCGGGGTTCCGGCGGCTCCGACCG 142
 QY 209 CCGG-----GGCGGCTACAGGCGCCACCGGAGCCTGAGAGTCGC 250
 Db 143 CGCGCGGAGCGGCTCTACGCGCGGAGGGCGGGCGCGGATCGGGAGCTTGAGGTCGC 202
 QY 251 TGACCGGGCGGACCTGCGGTGCTGCGGGCTCGCGGTCACAGCCGTCGACCGCT 310
 Db 203 TGAGCGGGCGGTGGCGAGCGGCTGTGTGGGCTCGCGGTCGAGCCGTCGCCCG-- 260
 QY 311 CGGACTCCGTGCGAAGAAATCGTCCGAGGTGATGAAACTCGCAAACTCCGTTGACGA 370
 Db 261 ----CTCCGTGCGCAAGAAATCGTCCGAGGTGATGAAACATTCGATTGATGA 316
 QY 371 AGCTCTCTGTTGAAGAGAAATCAGAAGAAGTTCTGTTCTACTTGAACGGGAGGTGAT 430
 Db 317 AGCCCTCTTGCTAAAGAGAAATCAGAAGAAGTTCTTCTTATTTGAAATGGACGGTGTAT 376
 QY 431 TTACCTAGTAGGAATCATGGGTTCTGGAAGAGTACTGTGGGAAGATTATGTCGAAGT 490
 Db 377 TTACCTAGTTGGAAATGGGTTCTGGAAGAGTACTGTGGGAAGATCATGTCGAAGT 436
 QY 491 CTTGGGTTATTCGTTCTTCATAGTCAGCAAGTTAGTGGAGCAAGCTGTTGGAATGCCATC 550
 Db 437 TTTGGGTTATTCGTTCTTTTGATAGTATAAATTTGGTCGAACAGCTGTGGCATGCCCTC 496
 QY 551 AGTTGCCCAATATTCAGGTCATAGTCAGGCTTCTTTCCGGGATAATGAGAGTAGTGT 610
 Db 497 AGTCGCTCAAAATTTTCAAGGTTTCATAGTGAAGCTTCTTTTGGGATAATGAGAGTAGTGT 556
 QY 611 CTTGAGAGATTGCTCTCATGCGAGCATAGTTAGTTGTTGCCACGGAGGTGGTCTGTAT 670
 Db 557 CTTGAGGATTGCTCTCATGAGCGATTAGTTGTTGCTACTGAGAGGTGGTCTGTAT 616
 QY 671 CCGACCAATTAACCTGGAGATATATGAAGAGGGGCGCTTATCTGTTGGTTAGATGTGCCCTT 730
 Db 617 CCGACCAATTAACTGGAATATACATGAAGAGG-----CCCT 653
 QY 731 GGAATGCTTGTCTAGGGGTTATGCTAAAGTGGGAAGTGGCTCTCGCTCTTCTTGACCA 790
 Db 654 GGACGCTCTGTCTAGGGGTTATGCTAAAGTGGGAAGTGGCTCCGCTCTTCTTAGATCA 713
 QY 791 ACCATCTGGTGTATCCGTACGCAATGGCTTTTCTAAGCTCA 831
 Db 714 ACCATCTGGTGTATCCATACAAATGGTGACTACTTGGCCCA 754

RESULT 4
 BT009514

LOCUS BT009514 1323 bp mRNA linear PLN 20-JUN-2003
 DEFINITION Triticum aestivum clone wrl.pk0099.b12:fis, full insert mRNA
 sequence.
 ACCESSION BT009514
 VERSION BT009514.1 GI:32129065
 KEYWORDS FLI CDNA.
 SOURCE Triticum aestivum (bread wheat)
 ORGANISM Triticum aestivum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Poideae; Triticeae; Triticum.
 1 (bases 1 to 1323)
 Tingey,S.V., Wolters,P., Powell,W., Dolan,M., Miao,G.-H.,
 Caraher,N.R., Hanafey,M.K. and Hainey,C.F.
 Direct Submission
 Submitted (20-JUN-2003) Crop Genetics, E. I. DuPont de Nemours and
 Company, 1 Innovation Way, P.O. Box 6104, Newark, DE 19714-6104,
 USA
 Location/Qualifiers
 1..1323
 /organism="Triticum aestivum"
 /mol_type="mRNA"
 /db_xref="taxon:4565"
 /clone="wrl.pk0099.b12:fis"

ORIGIN
 Query Match 30.4%; Score 364.4; DB 8; Length 1323;
 Best Local Similarity 71.9%; Pred. No. 2.4e-70;
 Matches 476; Conservative 0; Mismatches 186; Indels 0; Gaps 0;
 QY 333 TCCGAGGTCATGAAAACTCGCACAACTCCGTTGACGAAGCTCTCTGTTGAAGAGAAAA 392
 Db 170 TCGGAGGTACTGAGAAAGCCCACTATTCTGCTGATGAGGCTCTGCTACTAAAGCAAAA 229
 QY 393 TCAGAAGAAGTTCTGTTCTTACTTTGAACGGGAGTGTTATTTACTAGTAGGAATGATGGT 452
 Db 230 CGAGAGGACGTGCTCCCTTACCTGAATGACCGCTGTGTTATCTAGTTGGAATGATGGT 289
 QY 453 TCTGAAAAAGTACTGTGGGAGAGATTATGTTGAAGTCTTGGGTATTCTGTTCTTTGAT 512
 Db 290 TCCGCAAAACTACAGTTGGGAAGATAATAGCTGAAGTACTAGGCTATTTCATTCTTTGAC 349
 QY 513 AGTGCAAGTTAGTGAGCAAGCTCTTGGAAATGCCATCAGTTGCCCAATATTTCAAGGTC 572
 Db 350 AGTGATAGCTGTTGAGCAGCTCTGTGGCATACCGTCGCTGGCTGAGATTTTTCAGGTC 409
 QY 573 CATAGTAGAGCTTTCTTTCGGGATTAATGAGAGTAGTGTTTGAGAGATTTTGTCTCCATG 632
 Db 410 CACAGTGAAGCAATTTCTCAGAGATAACGAGAGTGAGTACTAAGGGATTTGTCTCAATG 469
 QY 633 CGACGATTAGTTGTTGCCACCGGAGGTGCTGTTATCCGACCAATTAACCTGGAGATAT 692
 Db 470 CACCGATTAAATTTGTGCAACAGGAGGTGCTGCGGTGATACGACCAATCAATTTGGAGTTAT 529
 QY 693 ATGAAGAGGGGCTATCTGTTTGGTTAGATGTGCCCTTGGATGCTCTTCTGCTAGGGGTATT 752
 Db 530 ATGAAGAAAGGACTCACTATTGTTGGTTAGATGTTCCATTGGACGCCCTTGCAGAGAGATT 589
 QY 753 GCTAAAGTGGGAATGTGCTCTCTGCTCTTCTTGGACCAACCATCTGCGGTGATCCGTACGCA 812
 Db 590 GCTCGGTTGGTACTCTGCGTCACGACCCCTCTCTGCATCAGGAATCTGGTGATCTTATGCA 649
 QY 813 ATGGCTCTTTCTTAAGCTCAGCATGCTGCACAGCAAGGGGTGATGCTTATCAATGCAATGCA 872
 Db 650 AAGGCTATGCCAAACTTACACACTTTTGAACAAAGAAATGGATTCATATGCTAATGCT 709
 QY 873 GATTAAGGGTTTCTCTGGAAGAGATTGTCATGTAAACCAAGGTCATGATGATCTCTTAAG 932
 Db 710 GATGCCGAGTTTCCCTTGAAAATATTGCAATTCAAACAAGGACATAATGATGTGAATGTA 769
 QY 933 CTGACACCTACTGATATTGCAATTGAATCAGTCACTTCAATAGATCGAGAGCTTCGTCATCGAG 992
 Db 770 CTTACACCAAGTGCCATCGCTATTGAGGCAATTTGCTAAAGATGGAGAGCTTTCTTACTGAG 829

```

QY      993 CA 994
Db      830 AA 831

RESULT 5
LOCUS   LESHKINP               1293 bp      mRNA      linear      PLN 16-JUN-1992
DEFINITION L.esculentum mRNA for shikimate kinase precursor.
ACCESSION X63560
VERSION   X63560.1 GI:19348
KEYWORDS  shikimate kinase.
SOURCE    Lycopersicon esculentum (tomato)
ORGANISM  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
REFERENCE
AUTHORS  Schmid,J., Schaller,A., Leibinger,U., Boll,W. and Amrhein,N.
TITLE    The in-vitro synthesized tomato shikimate kinase precursor is
            enzymatically active and is imported and processed to the mature
            enzyme by chloroplasts
JOURNAL  Plant J. 2 (3), 375-383 (1992)
MEDLINE  93272054
PUBMED   1338949
REFERENCE
AUTHORS  Schaller,A.
TITLE    Direct Submission
JOURNAL  Submitted (27-DEC-1991) A. Schaller, Federal Institute of
            Technology, Institute of Plant Sciences, Sonnegstrasse 5, 8092
            Zuerich, SWITZERLAND
FEATURES
Source    1. .1293
            /organism="Lycopersicon esculentum"
            /mol_type="mRNA"
            /strain="UC92b"
            /db_xref="taxon:4081"
            123. .1025
            /codon_start=1
            /product="shikimate kinase precursor"
            /protein_id="CAA45121.1"
            /db_xref="GI:19349"
            /db_xref="GOA:Q00497"
            /translation="MEARVQSILQSLWSNDSKVRKPSGLLRFSEKWNKPRHRVVV
            SCHLOPKAHSRRVQLKVSQPONQASVLESQFSASIDEIETLKNKAEVEEYL
            DRCVYLGMCGKTTVGRILAEILGYSFFDCDLRIEQAVGIVTAIEIFELKGSFF
            RNETEVLHLKSLMHLVSVSTGGAVVRPINRHHKIGISVLDVPLEALAKRITTEG
            TKSRLHESGSDVDTTLKRLTLLMETRGYANASARVSLNIALKREKDVCHITP
            AEITLEVLQIENFLKTKQSVVL"
            123. .320
            transit_peptide
            mat_peptide 321..1022
            /product="shikimate kinase"
            /EC_number="2.7.1.71"

ORIGIN
Query Match      20.2%; Score 242.4; DB 8; Length 1293;
Best Local Similarity 62.7%; Pred. No. 3.4e-43;
Matches 394; Conservative 0; Mismatches 231; Indels 3; Gaps 1;

QY      360 TCGTGTGACGAAGCTCTCTCTGTTGAAGAGAAAATCAGAGAAGTTCTGTCTACTTTGAAC 419
Db      372 TCAATTGATGAATCGAGACATTAAGAATAAGCAGAGAGAGTTGAAGAATATCTAGAT 431

QY      420 GGGAGGTGTATTACTAGTAGAATGATGGGTTCTGGAAAAAGTAGTCTGTGGGGAAGATT 479
Db      432 GCACGATGTGTATACCTTGTGGAATGATGGGCTGCGGAAAAACAACCTGTGGGCCGGATT 491

QY      480 ATGCTCTGAAGTCTTGGGTTATTCTGTTCTTTGATGACAGAGTTAGTGGGCAAGCTGTT 539
Db      492 TTGGCAGAAACACTGGGATATTCCTTTTGTGACTGTGACAGGCTGATAGACAGGCTGTT 551

QY      540 GGAATGCCATCAGTTTGCCCAAAATATTCAAGGTCCCTAGTAGAAGCCCTTTCTTCGGGATAAT 599
Db      552 GGTGGAAATTACTGTAGCTGAATCTTTGAGCTTCGCGGAGAGAGCTTCTTCAGGGATAAT 611

QY      600 GAGAGTAGTGTCTTGAGAGATTGTCCTCCATCGGACGATTAGTTGTTGCCACCGAGGT 659
Db      612 GAGACTGAGGTATTGCACAAACTATCTTTGATGCAATCGGCTTGTGTTTCAACGGGTGA 671

QY      660 GGTGCTGTTATCCGACCAATTAACCTGGAGATATATGAAGAGGGGCTATCTGTTTGGTTA 719
Db      672 GGTGAGTGTGTCGTCCTCCATTAATTTGGAGCAATATGCACAGGATATTAGTTGTTGTTA 731

QY      720 GATGTCCTTCGATGCTCTTCTAGGCGTATTGCTAAAGTGGGAACCTCCTCTCTCTCT 779
Db      732 GATGTTCTTTAGAAGCTTTGGCCCAAGAGAATTACACGGAAGGAACATAAATCTCGCCCC 791

QY      780 CTTCTGGACCAACCATCTGGTGATCGGTACGCAATGGCCTTTTCTAAGCTCAGCATGCTT 839
Db      792 CTTTACATGAAGAATCAGGAGACCTTTATGATACGACTTTTGAAGCGGTAACTACTTTA 851

QY      840 GCACAGCAAAAGGGGTGATGCTTTATGCAATGCAAGATGTAAGGGTTTCTCTGGAAGAGATT 899
Db      852 ATGGAGACGAGGGGTGAAAATTTATGCCAATGCAGTGCAGGGTTTCTACTAGAAAATATT 911

QY      900 GCATGTAAACAAGGTCATGATGATGCTCTCTAAGCTGACACCTACTGATATTGCAATTGAG 959
Db      912 GCATTGA---AAAGGGAAGAAAGATGCTGTCTATATTACACCTGCTGAAATATTACTTAGAG 968

QY      960 TCACCTTCATAGATCGAGAGCTTCGTCA 987
Db      969 GTTCTTATACAAATTGAGAACTTCTTAA 996

RESULT 6
LOCUS   BT012939               1303 bp      mRNA      linear      PLN 11-MAY-2004
DEFINITION Lycopersicon esculentum clone 114107R, mRNA sequence.
ACCESSION BT012939
VERSION   BT012939.1 GI:47104354
KEYWORDS  FLI_CDNA.
SOURCE    Lycopersicon esculentum (tomato)
ORGANISM  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
REFERENCE
AUTHORS  Kirkness,E.F., Wang,W. and Vazeille,A.
TITLE    Direct Submission
JOURNAL  Submitted (11-MAY-2004) The Institute for Genomic Research, 9712
            Medical Center Drive, Rockville, MD 20850, USA
FEATURES
Source    1. .1303
            /organism="Lycopersicon esculentum"
            /mol_type="mRNA"
            /db_xref="taxon:4081"
            /clone="114107R"
            /tissue_type="mature green fruit"
            /note="TWGE47"

ORIGIN
Query Match      20.2%; Score 242.4; DB 8; Length 1303;
Best Local Similarity 62.7%; Pred. No. 3.4e-43;
Matches 394; Conservative 0; Mismatches 231; Indels 3; Gaps 1;

QY      360 TCGTGTGACGAAGCTCTCTCTGTTGAAGAGAAAATCAGAGAAGTTCTGTCTACTTTGAAC 419
Db      344 TCAATTGATGAATCGAGACATTAAGAATAAGCAGAGAGAGTTGAAGAATATCTAGAT 403

QY      420 GGGAGGTGTATTACTAGTAGAATGATGGTTCCTGAAAAAGTACTGTGGGGAAGATT 479
Db      404 GCACGATGTGTATACCTTGTGGAATGATGGGCTGCGGAAAAACAACCTGTGGGCCGGATT 463

QY      480 ATGCTCTGAAGTCTTGGGTTATTCTGTTCTTTGATGACAGAGTTAGTGGGCAAGCTGTT 539

```


ACCESSION AK117791.1 GI:26450650
 VERSION FLI cDNA; CAP trapper.
 KEYWORDS Arabidopsis thaliana (thale cress)
 SOURCE Arabidopsis thaliana
 ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

1 Seki, M., Iida, K., Satou, M., Sakurai, T., Akiyama, K., Ishida, J., Nakajima, M., Enju, A., Kamiya, A., Narusaka, M., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.
 Arabidopsis thaliana full-length cDNA
 Published Only in Database (2002)

2 (bases 1 to 1294)
 Seki, M., Iida, K., Satou, M., Sakurai, T., Akiyama, K., Ishida, J., Nakajima, M., Enju, A., Kamiya, A., Narusaka, M., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.
 Direct Submission
 Submitted (25-NOV-2002) Motoaki Seki, RIKEN Genomic Sciences Center; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: mseki@gsc.riken.go.jp, URL: http://pfjweb.gsc.riken.go.jp, Tel: 81-45-503-9625, Fax: 81-45-503-9586)

An Arabidopsis full-length cDNA library was constructed essentially as reported previously (Seki et al. (1998) Plant J. 15:707-720; Seki et al. (2002) Science 296:141-145). cDNA cleaved with BamHI and XhoI was ligated to modified Lambda FLC-1-E vector (Carninci et al. (2001) Genomics 77:79-90) digested with BamHI and SalI. This clone is in a modified pBluescript vector.
 Please visit our web site (http://pfjweb.gsc.riken.go.jp/) for further details.

FEATURES
 Location/Qualifiers
 1..1294
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /db_xref="taxon:3702"
 /chromosome="4"
 /clone="RAFL17-48-C24"
 /ecotype="Columbia"
 /note="common name: thale cress"
 1..1294
 /gene="At4g39540/F23K16_170"
 228..1130
 /gene="At4g39540/F23K16_170"
 /codon_start=1
 /product="putative shikimate kinase"
 /protein_id="BAC42436.1"
 /db_xref="GI:26450651"
 /translation="MEAAVQRFQYSSWDLNRFEGKPGSLRYNTQRIKEDRRFVV ALTDKRRDRLRSVSKNSALLTSPDFDEQOILKKKAEVXPYLNRSMTY LVGMGSGKTTVGKINARSLGYTFDCDITLIEQAMKGTVAEIEFHFEGSVFREKETE ALKKLSIMYHVVSVGGGAVIRPINWYMKHGISIWLDVPLEALHRAAVGTGSRP LLDHDSGDTYTAALNRLSTIMDARGEATTKASARVSLNITLKLGYRSVSLTFAEI AIEAFQVQSYLEKEDGMARDPL"

Query Match 16.1%; Score 193; DB 8; Length 1294;
 Best Local Similarity 59.0%; Pred. No. 3.3e-32;
 Matches 369; Conservative 0; Mismatches 250; Indels 6; Gaps 2;

QY 363 GTTGACGAGCTCTCCTGTTGAAGAGAAAATCAGAGAAAGTTCTGTCTTCTACTTGAACGGG 422
 Db 462 GATGAAGACAACAGATTTGAAGAAAAAGCTGAAGAGTTAAACCGTATTAAATGA 521

QY 423 AGGTGATTATTACCTAGTAGAATGATGGTTCGGAAGAAAGTACTCTGGGGAAGATTATG 482
 Db 522 CGATCATGATGATCTTGTGTGATGATGGTTCGGGAAAACGACTAGGGAAGATTATG 581

QY 483 TCTGAAGTCTTGGGTTATCGTCTTTGATAGTGAAGTACTAGTGAAGCAAGCTGTTGA 542
 Db 582 GCAAGATCGCTTGGTTATACATTTCTTTGATTTGTGACACTTTGTATCGAGCGGCTATGAAG 641

QY 543 ATGCCATCAGTTGCCCAAAATATTCAAGGTCATAGTAGGAAGCCCTTTCTTGGGATAATGAG 602
 Db 642 GGAACTTCTGTAGCTCAGATATTTTGAGCATTTTCGGGTGAGAGTGTCTTTAGAGAAAAAGAG 701

QY 603 AGTAGTGTCTTGAGAGATTGTCTCCTCATG---CGACGATTAGTTGTGTGCCACCGAGGT 659
 Db 702 ACTGAAGCGTTAAAGAAACTCTCTTTGATGTATACCACCAAGTTGTGTGTTTCAACCGGGGA 761

QY 660 GGTGCTGTTTATCCGACCAATTAACTGGAGATATATGAAGAGGGGCGCTATCTGTTTGGTTA 719
 Db 762 GGGGCAATTATAGACCCATCAATTGGAAGTACATCATTAAGGTAATTAGTATTGGCTT 821

QY 720 GATGTGCCCTTGAGTCTCTTGTAGGCGTATTGTCTAAAGTGGGAACTGCCTCTCGTCT 779
 Db 822 GATGTACTCTAGAAGCCTTAGCGCATAGATAGTGTCTGTAGGAACCTGTTTCAAGACCA 881

QY 780 CTTCTG--GACCAACCATCTGGTGTATCGTAGGCAATGGCCTTTTCTAAGCTCAGCATG 836
 Db 882 TTGTACATGATGATGAGTCAGGGGACACATACACAGCGGCTTTAAACCGCTCTTCAACG 941

QY 837 CTTGCACAGCAAAAGGGGTGATGCTTATGCAAAATGCAGATGTAAAGGTTTCTCTGGAAGAG 896
 Db 942 ATTGGGATGCAGTGTGTGAGCATACATTAAGCCAGCGCAGAGATTTCTTGGAGAT 1001

QY 897 ATTGCATGTAAACAAGGTCATGATGATGCTCTCTAAGCTGCACACCTACTGATATTGCAATT 956
 Db 1002 ATTACTTTCAAGCTCGTTTATAGAAGTGTCTCAGATCTTTACACCAGCTGAAATCGCAATT 1061

QY 957 GAGTCACTTCATAAGATCGAGAGCT 981
 Db 1062 GAGGCGCTTTGAGCAAGTTCAGAGCT 1086

RESULT 11
 AY596190
 LOCUS Cucumis sativus shikimate kinase mRNA, partial cds.
 DEFINITION
 ACCESSION AY596190
 VERSION AY596190.1 GI:46849855
 KEYWORDS Cucumis sativus (cucumber)
 ORGANISM Cucumis sativus
 Cucumis sativus
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Cucurbitales; Cucurbitaceae; Cucumis.
 1 (bases 1 to 1016)
 Kim, M.S., Yang, K.Y., Kim, Y.C. and Cho, B.H.
 A shikimate kinase and an acidic peroxidase gene expression were primed by a rhizobacterial colonization upon challenge-inoculation with *Corynebacteria cassicola* in cucumber leaves
 Unpublished
 JOURNAL
 REFERENCE 2 (bases 1 to 1016)
 Kim, M.S., Yang, K.Y., Kim, Y.C. and Cho, B.H.
 Direct Submission
 Submitted (09-APR-2004) Applied Plant Science Division, Chonnam National University, Yongsong Dong 300, Bukgu, Gwangju 500-757, Korea
 Location/Qualifiers
 1..1016
 /organism="Cucumis sativus"
 /mol_type="mRNA"
 /cultivar="Baikseung 3"
 /db_xref="taxon:3659"
 1..474
 /codon_start=1
 /product="shikimate kinase"
 /protein_id="AA02351.1"
 /db_xref="GI:46849856"
 /translation="GISVAEVFKVYGEDPFRERTEALRKLRLMSQFVISTGGAVTR SINWYMKHGISVWLDVPLEALVKRISAVGNSRPLHSDNSDAISKTLVRLSTLEE RGEAYANAERKVSCEKIAAKUGTKDVSNVTPMAIAIEALEELETFLKREDGYCAF"

FEATURES
 source
 CDS

ORIGIN

Query Match	15.5%;	Score 185.8;	DB 8;	Length 1016;
Best Local Similarity	61.8%;	Pred. No. 1.3e-30;	Indels	0; Gaps 0;
Matches 295;	Conservative	0;	Mismatches 182;	
QY	549	TCAGTTGGCCCAATATTCGAAGTCCATAGTGAAGCCTTCTTTCCGGGATAATCAGAGTAGT	608	
Db	7	TCTGTAGCCGAAGTTTTCGAAGTCTATGCGGAGGACTTCTTTAGAGAAAGGAGCGGAA	66	
QY	609	GTCCTGAGAGATTGTCCTCCATCGGACGATTAGTTGTCGCCCGGAGTGGTCTGTT	668	
Db	67	GCAATTGAGGAAGTTATCTTTGATGCGCCCAATTTGTTATTTCACAGCGGAGTGCTGTA	126	
QY	669	ATCCGACCAATTAACGGAGATATATGAAGAGGGGCTATCTGTTGGTTAGATGCGCC	728	
Db	127	ACGAGATCGATAACTGGAAATATATGCAATAAGGCATCAGTGTGTTGGTATGTCCT	186	
QY	729	TTGGATGCTCTTGCTAGGCGTATTGCTAAAGTGGGAACCTGCTGCTCTTCTTCGGAC	788	
Db	187	TTGGAAGCCTTGTTGAAGAGAATTCAGCTGTAGGAACCAATTCCTGCTCCCTTTTGCAAT	246	
QY	789	CAACCATCTGGTATCCGTACGCAATGGCTTTTCTAAGCTCAGCATGCTTGCACAGCA	848	
Db	247	CATGATTCAAATGATCATACTCAAGACCCCTCGTCCGTTTATCTACTCTTCGGAAGAG	306	
QY	849	AGGGGTGATGCTTATGCAAAATGCAGATGAAGGTTTCTCTGGAAGAGATTGCATGTAA	908	
Db	307	AGGGGCAAGCATATGCCAACGCTGAAGTCAAAAGTTTCTCGAATAAATTCGAGCCAA	366	
QY	909	CAAGGTCATGATGATGCTCTTAAGCTGCACACCTTACTGATATTCGAATTGAGTCAC	968	
Db	367	CTGGGTACTAAAGACGTATCGAATGTCACGCCTATGCGCATAGCAATCGAGGCACTCGAA	426	
QY	969	AGATCGAGGCTTCTGTCATCGAGCACCTGCTGATAGTTTCAGTCAGCACCGCA	1025	
Db	427	GAATCGAGACCTTTTGAAGAGAGAAGATGGTTATTGTGCATTTTGTATGACCA	483	

AK067948 2127 bp mRNA linear PLN 24-JUL-2003
Oryza sativa (japonica cultivar-group) cDNA clone:J013123F09, full insert sequence.

AK067948 1 GI:32977966
FLI_CDNA: CAP trapper.

Oryza sativa (japonica cultivar-group)
Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.

1 The Rice Full-length cDNA Consortium, National Institute of Agricultural Sciences Rice Full-length cDNA Project Team: Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Nami, T., Ohneda, E., Yabagi, W., Suzuki, K., Li, C., Ohtsuka, K., Shishiki, T., Foundation of Advancement of International Science Genome Sequencing & Analysis Group: Ohtsuka, K., Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda, Y., Kuroseki, T., Kodama, T., Masuda, H., Kobayashi, M., Xie, Q., Lu, M., Narioka, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Niikura, J., Ikeda, R., Ishibiki, J., Kawamata, M., Yoshimura, A., Miura, J., Kusumegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., RIKEN, Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashidume, W., Hayatsu, N., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kondo, S., Konno, H., Miyazaki, A., Osato, N., Ota, Y., Saito, R., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Yoshino, M., and Hayashizaki, Y.

Collection, mapping, and annotation of over 28,000 cDNA clones from japonica rice

JAPONICA RICE
Science 301 (5631), 376-379 (2003)
22752273
12869764

2 (bases 1 to 2127)
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K., Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Hotta, I., Iida, J., Itoh, Y., Ikeda, R., Imamura, K., Imotani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I., Kanagawa, S., Kato, H., Kawagashira, N., Kawai, J., Kobayashi, M., Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M., Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koyama, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numasaka, R., Ohno, M., Osato, N., Ota, Y., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Takami-Takeda, Y., Tagawa, A., Takahashi, F., Tanaka, T., Tomaru, A., Toya, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yabagi, W., Yanada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S. and Yoshimura, A.

Direct Submission
Submitted (05-DEC-2001) Shoshi Kikuchi, National Institute of Agricultural Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression: 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan [E-mail:skikuchi@nias.affrc.go.jp, Tel:81-29-838-7007, Fax:81-29-838-7007]

This clone is one of the 28K full-length cDNA clones from japonica rice.

URL : http://cdna01.dna.affrc.go.jp/cDNA/
NIAS Rice Full-length cDNA Project Team: Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Nami, T., Ohneda, E., Yabagi, W., Suzuki, K., Li, C., Ohtsuka, K., Shishiki, T. and Yamamoto, M.

FAIS Genome Sequencing & Analysis Group: Ohtsuka, Y., Iida, Y., Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M., Kodama, T., Kuroseki, T., Kusumegi, T., Lu, M., Masuda, H., Miura, J., Mizuno, K., Narioka, R., Niikura, J., Oka, M., Ryu, R., Sugano, S., Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S., Yoshimura, A., Matsubara, K. and Murakami, K.

Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Iida, J., Imamura, K., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kanagawa, S., Kato, H., Kawai, J., Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koyama, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numasaka, R., Ohno, M., Osato, N., Ota, Y., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Takami-Takeda, Y., Tagawa, A., Takahashi, F., Tanaka, T., Tomaru, A., Toya, T., Waki, K., Yasunishi, A. and Hayashizaki, Y.

Location/Qualifiers
1. 2127
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="J013123F09"

FEATURES
source

ORIGIN
Query Match 13.7%; Score 164.4; DB 8; Length 2127;
Best Local Similarity 69.8%; Pred. No. 8.3e-26;
Matches 222; Conservative 0; Mismatches 96; Indels 0; Gaps 0;

QY 677 AATTAACTGGAGATATATGAGAGGGGCTATCTCTTTGGTTAGATGTCCTTGGATGC 736
DB 1584 AATTTCGAGGAGTTATCATGAAGAAGGGTCAACAATCTGTTAGATGTTCCATTGATGC 1643

QY 737 TCTGTAGCGGTAATGCTAAAGTGGGAAGTGGCTCTCGTCTCTTCTGGAACAACATC 796
Db 1644 ACTTGAAGACGAATATGCTGAGTAGGAATGCTTCTCGGCCCTCTTGATCAAGAAATC 1703
QY 797 TGGTGATCCGTAGCGAATGCGCTTTCTAAGCTCAGCATGCTTGCACAGCAAGGGGTGA 856
Db 1704 TGGTGATCCCTTATGCAAGAGCTTATGCAAACTGACTGACATTTTGAACAAGAAATGA 1763
QY 857 TGCTTATGCAATGCAAGATGTAAGGGTTTCTCTGGAAGAGATGTCATGTAACAAGGTCA 916
Db 1764 CTCATATGCTAATGCTGATGCCAGGGTTTCACTCGAACATATGCAAGTAAGACGACA 1823
QY 917 TGATGATGCTCTTAAGCTGACACCTACTGATATGCAATGCAATGAGTCACTTCATAAGATGA 976
Db 1824 TAGTAATGTGACGACACTTACACCAAGTGCCATTTGCTATTGAGGCAATTGCTAAAGATGA 1883
QY 977 GAGCTTCGTCATCGACA 994
Db 1884 GAGCTTCTTACTGAGAA 1901

RESULT 13
AX315616
LOCUS AX315616 309 bp DNA linear PAT 14-DEC-2001
DEFINITION Sequence 8601 from Patent WO0190366.
ACCESSION AX315616
VERSION AX315616.1 GI:17899043
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
REFERENCE
AUTHORS Leach,M.D. and Shinkets,R.A.
TITLE Human polynucleotides and polypeptides encoded thereby
JOURNAL Patent: WO 0190366-A 8601 23-NOV-2001;
Curagen Corporation (US)
FEATURES
Location/Qualifiers
1..309
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 13.1%; Score 157; DB 6; Length 309;
Best Local Similarity 69.3%; Pred. No. 2.6e-24;
Matches 214; Conservative 0; Mismatches 95; Indels 0; Gaps 0;
QY 681 AACTGAGATATATGAAGAGGGCCCTATCTGTTGTTAGATGTGCCCTTGGATGCTTT 740
Db 1 AATTGGAGTTACATGAAGAAAGGGCTGACCGTATGTTAGATGTCCCACTGGATGCATT 60
QY 741 GCTAGCGCTATTGCTAAAGTGGGAAGTGGCTCTCGTCTCTTCTGACCAACCATCTGCT 800
Db 61 GCAAGAAGAATGCTGCTGTAGGAACCGGCTCTGACCACTCTTGGATCAGGAATCCGGT 120
QY 801 GATCCGTAGCAATGCGCTTTTCTAAGCTCAGCATGTTGCGACAGCAAAAGGGGTGATGCT 860
Db 121 GATCCTTATGCAAGAGCTTATGCAAACTTACGTCACTTTTTCGACGAAGATGACTCG 180
QY 861 TATGCAATGCGAGATGTAAGGGTTTCTCTGGAAGAGATGTCATGTAAACAAGGTATGAT 920
Db 181 TATGCTAATGCTGATGCCAGAGTTTCACTTGAACATATTGCAATTAACAACAGGCCATAT 240
QY 921 GATGCTCTTAAGCTGACACCTACTGATATTGCAATTTGCAATGAGTCACTTCAATGATCGAGC 980
Db 241 GATGCTCACTATACCTACCTAGTACCATCGCAATTGAGGCATTTGCTAAAGATGGGAAGT 300
QY 981 TTGCTCATC 989
Db 301 TTTCCTACC 309

RESULT 14
AP004089/c
LOCUS
DEFINITION
QY 111777 bp DNA linear HTG 21-MAR-2002
Db Oryza sativa (japonica cultivar-group) chromosome 2 clone
AP004089.1 GI:15281368
ACCESSION HTG; HTGS PHASE2.
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES
source
ORIGIN
Query Match 10.0%; Score 120.4; DB 2; Length 111777;
Best Local Similarity 86.4%; Pred. No. 1.1e-15;
Matches 133; Conservative 0; Mismatches 21; Indels 0; Gaps 0;
QY 678 ATTAAGTGGAGATATGAGAGGGCCCTACTGTTGGTTAGATGTCCTTGGACCATCT 737
Db 64333 ATTATCAGGAATATACATGAAGAGGGCCCTATCTGTTGGTGGATGTGCGCTTGGACGCT 64274
QY 738 CTTGCTAGCGCTATTGCTTAAAGTGGGAAGTGGCTCTCGTCTCTTCTTGACCAACCATCT 797
Db 64273 CTTGCTAGCGCTATTGCTTAAAGTGGGACTGCTCCCGCTCTCTTAGATCAACCATCT 64214
QY 798 GGTGATCCGTACGCAATGCGCTTTTCTAAGCTCA 831
Db 64213 GGTGATCCATACACAATGCTGACTACTTGGGCCCA 64180
RESULT 15
AP004774
LOCUS
DEFINITION
QY 154796 bp DNA linear PLN 15-APR-2004
Db Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 2,
AP004774 PAC clone:P0431B06.
ACCESSION AP004774
VERSION AP004774.3 GI:46390195
KEYWORDS

SOURCE ORGANISM	Oryza sativa (japonica cultivar-group) Oryza sativa (japonica cultivar-group) Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.	gene	complement (4342. .4808) /gene="P0431B06.2"
REFERENCE		mRNA	complement (4342. .4808) /gene="P0431B06.2"
AUTHORS	1	CDS	note="supported by full-length cDNA (s) : AK058832" complement (4442. .4759) /gene="P0431B06.2"
TITLE	Sasaki, T., Matsumoto, T. and Yamamoto, K.		note="contains full-length cDNA (s) : AK058832" /codon_start=1 /product="unknown protein" /protein_id="BAD15627.1" /db_xref="GI:46390196" /translation="MKVAVLVGRGPRREREREGCVRRRGGRRAPDPARGRVVGR RIGVGVRRLLDPASRPPEGAAVRGGARRRPRAVAHTCRCLLPSPSHMPATASAG IEE"
JOURNAL	Submitted (20-FEB-2002) Takuji Sasaki, National Institute of Agrobiological Sciences, Rice Genome Research Program; Kannonnai	gene	complement (6499. .7930) /gene="P0431B06.3"
REFERENCE	2-1-2, Tsukuba, Ibaraki 305-8602, Japan (E-mail:tsasaki@ias.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/, Tel.81-298-38-7441, Fax:81-298-38-7468)	mRNA	complement (join(<6499. .6558,6932. .7039,7135. .7255, 7869. >7930)) /gene="P0431B06.3"
AUTHORS	Direct Submission		note="start and end point are not identified" complement (join(6499. .6558,6932. .7039,7135. .7255, 7869. .7930)) /gene="P0431B06.3"
JOURNAL	Submitted (20-FEB-2002) Takuji Sasaki, National Institute of Agrobiological Sciences, Rice Genome Research Program; Kannonnai	CDS	note="contains EST (s) : AU172742 (E20094), C73662 (E20094)" /codon_start=1 /product="dolichol-phosphate mannosyltransferase-like" /protein_id="BAD15628.1" /db_xref="GI:46390197" /translation="MVWCICIERQLVARDPRSLSSKHLVSQMKHIFKIIAMLVAVSAI WIALETATVPRSVTWLLPIVLVALGCYGLFMVGFGLMFPFPTCPQEAULLQDDILEA KEFLSKGVDVGSSE"
COMMENT	On Apr 14, 2004 this sequence version replaced gi:33468820. Genes were predicted from the integrated results of the following: GENSCAN (http://ccr-081.mit.edu/GENSCAN.html), FGENESH (http://www.softberry.com/), GeneMark.hmm (http://opal.biology.gatech.edu/GeneMark/), GlimmerM (http://www.tigr.org/tcdb/glimmer/glmr form.html), RiceHMM (http://rgp.dna.affrc.go.jp/RiceHMM/), _SplicePredictor (http://bioinformatics.lasstate.edu/cgi-bin/sp.cgi), sim4 (http://globin.cse.psu.edu/html/docs/sim4.html), gap2 (http://www.tigr.org/software/glimmer/), BLASTN and BLASTX. The genomic sequence was searched against NCBI NonRedundant Protein database, nr (ftp://ncbi.nlm.nih.gov/blast/db) and the cDNA sequence database at RGP or DDBJ. Protein homologues of the coding regions were searched against NCBI NonRedundant Protein database with BLASTP. ESTs represent the identified cDNA sequences using BLASTN with the corresponding DDBJ accession no. and RGP clone ID. Full-length cDNAs represent the identified cDNA sequences using BLASTN with the corresponding DDBJ accession no. A gene with identity or significant homology to a protein is classified based on the protein name to indicate the homology level such as same name, 'putative-' and '-like protein'. A gene without significant homology to any protein but with full-length cDNA or EST homology (covering almost the entire length of partial sequence) is classified as an 'unknown' protein. A gene predicted by two or more gene prediction programs is classified as a 'hypothetical' protein according to IRGSP standard. A gene predicted by a single gene prediction program is also classified as a probable 'hypothetical' protein and is included as a miscellaneous feature of the sequence. The orientation of the sequence is from T7 to SP6 of the PAC clone. This sequence of P0431B06 clone has an overlap with OSJNBa0078N11 (DDBJ: AP005848) clone at 5' end and with OJL288 G09 (DDBJ: AP004119) at 3' end. Detailed information on overlap and assembly quality together with annotation of this entry is available at http://rgp.dna.affrc.go.jp/GenomeSeq.html. Location/Qualifiers	gene	join(9079. .9119,9183. .9349,9987. .10000) /gene="P0431B06.4"
		misc_feature	note="hypothetical ORF /note="predicted by GENSCAN this category is not included in IRGSP standard" complement (10085. .21473) /gene="P0431B06.5"
		gene	complement (join(10085. .10456,10633. .10695,10796. .10912, 11113. .11202,11457. .11544,12022. .12091,12505. .12583, 13791. .13835,14487. .14573,14661. .14744,14864. .15043, 15122. .15262,15433. .15541,15971. .16026,16437. .16484, 16581. .16719,17270. .17377,17698. .17724,17885. .17977, 18397. .18476,18611. .18700,18846. .18936,19063. .19354, 21171. .21473)) /gene="P0431B06.5"
		mRNA	note="supported by full-length cDNA (s) : AK119455" complement (join(10325. .10456,10633. .10695,10796. .10912, 11113. .11202,11457. .11544,12022. .12091,12505. .12583, 13791. .13835,14487. .14573,14661. .14744,14864. .15043, 15122. .15262,15433. .15541,15971. .16026,16437. .16484, 16581. .16719,17270. .17377,17698. .17724,17885. .17977, 18397. .18476,18611. .18688)) /gene="P0431B06.5"
		CDS	note="contains EST (s) : AU069794 (E3332) contains full-length cDNA (s) : AK119455" /codon_start=1 /product="putative SEC6" /protein_id="BAD15629.1" /db_xref="GI:46390198" /translation="MMSISVBAARASDLSDDKELIRYERLTALDGRERFALAAAS HKKEVGRUREYFVDVRSWETFERKALWGHVTFNFRKLSQTSQTILRAIRLVVEMQBIID QQVAEEAAEAGAGAMATIANQRRTANRKGGAATSPRSTQKSKVGKQYKCKEY IRKAVEARFDKLTELVEFSDIMEALEAKAGIDELGDIYDYVAPCFPRYEIFOLMV NLKATFERFQMLRLISDRANDIPNINILKVTGVVVKYQENLIGLVDESIAOVCSGSA LDPLMNVYERMQATTKWYSNILEADKTQPKSTEDCKLYTPRAVDLFRILITEOVQI VTENSTDVMLYRALAVIQVWMLDFQAAERQRLERPSASVGLSICALINNLRCYELS SELSSSTLLEAFQNAEQVNFEDCKGFLEVAKEAVLQTVSVIFDPGQDLILKYQ KWMMEGMVTEYIVATFDYFGDVKQYIEERSFRFVEACLEQTIIVVYDHLTLQKGI KEETIERMLDEVLMDPFREHINVTKETRVRILADURELASAESLDSFTLIYTNIL
FEATURES	source		
	1. .154796 /organism="Oryza sativa (japonica cultivar-group)" /mol_type="Genomic DNA" /cultivar="Nipponbare" /db_xref="taxon:39947" /chromosomes="2" /clone="P0431B06" 747. .3402 /gene="P0431B06.1" join(<747. .1033,1549. .1595,1683. .1737,1842. .1914, 2001. .2088,2218. .2300,2401. .2421,2445. .2821,2969. .3033, 3141. .>3402) /gene="P0431B06.1" note="supported by full-length cDNA (s) : AK066687" 747. .3402 /gene="P0431B06.1" note="contains full-length cDNA (s) : AK066687 non-coding transcript probably inactive due to 3'UTR too long in CDS"	gene	
		mRNA	
		misc_feature	

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 25, 2005, 12:54:01 ; Search time 726 Seconds
(without alignments)
9784.696 Million cell updates/sec

Title: US-10-660-226-9

Perfect score: 1200

Sequence: 1 ccgccaccagctacctgcc.....aaaaaaaaaaaaaaaaaaaa 1200

Scoring table: IDENTITY NUC

Gapop 10_0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_16Dec04:*

- 1: Geneseqn1980s:*
- 2: Geneseqn1990s:*
- 3: Geneseqn2000s:*
- 4: Geneseqn2001as:*
- 5: Geneseqn2001bs:*
- 6: Geneseqn2002as:*
- 7: Geneseqn2002bs:*
- 8: Geneseqn2003as:*
- 9: Geneseqn2003bs:*
- 10: Geneseqn2003cs:*
- 11: Geneseqn2003ds:*
- 12: Geneseqn2004as:*
- 13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	239.4	20.0	912	3 AAC50023	Aac50023 Arabidops
2	236.8	19.7	1365	3 AAC50026	Aac50026 Arabidops
3	236.8	19.7	1491	3 AAC50029	Aac50029 Arabidops
4	236.2	19.7	1363	3 AAC36026	Aac36026 Arabidops
5	233.6	19.5	637	10 ADK52674	Adk52674 Plant DNA
6	227	18.9	1484	3 AAC40184	Aac40184 Arabidops
7	188.4	15.7	1095	3 AAC47100	Aac47100 Arabidops
8	181	15.1	1093	3 AAC36746	Aac36746 Arabidops
9	166.6	13.9	696	3 AAC41622	Aac41622 Arabidops
10	163	13.6	660	10 ADK58257	Adk58257 Plant DNA
11	157	13.1	309	6 ABN79354	Abn79354 Human kin
12	138.4	11.5	329	10 ADK58255	Adk58255 Plant DNA
13	135.2	11.3	573	13 ACN52580	Acn52580 Cotton an
14	130	10.8	1010	13 ADK60835	Adk60835 Cotton cd
15	125.8	10.5	614	13 ADK60836	Adk60836 Cotton cd
16	113.2	9.4	226	7 ADK67657	Adk67657 Corn seed
17	94.6	7.9	537	10 ADK58256	Adk58256 Plant DNA
18	85.4	7.1	570	13 ADS47877	Ads47877 Bacterial
19	82.6	6.9	255	7 ADK66821	Adk66821 Corn seed
20	79.4	6.6	402	13 ADT44359	Adt44359 Bacterial

21	74	6.2	564	13 ADT45202	Adt45202 Bacterial
22	66	5.5	2000	8 ADA71938	Ada71938 Rice gene
23	65.6	5.5	480	10 ADE82174	Ade82174 Arabidops
24	60.2	5.0	1479	13 ADS48191	Ads48191 Bacterial
25	59.2	4.9	507	13 ADS45422	Ads45422 Bacterial
26	58.6	4.9	498	13 ADS57732	Ads57732 Bacterial
27	56.2	4.7	218	12 ADQ17783	Adq17783 Human col
28	56.2	4.7	345	4 AAH33454	Aah33454 Human col
29	56	4.7	457	13 ADR60837	Adr60837 Cotton cd
30	56	4.7	457	13 ACN46460	Acn46460 Cotton pr
31	55.4	4.6	528	13 ADT46243	Adt46243 Bacterial
32	55.4	4.6	1682	5 ADM19284	Adm19284 Novel hum
33	55.4	4.6	1692	12 ADQ24979	Adq24979 Human sof
34	55.2	4.6	742	2 AAV00437	Aav00437 Clone H90
35	55	4.6	405	4 AA184799	Aai84799 Human pol
36	55	4.6	603	9 ADA30736	Ada30736 DNA encod
37	55	4.6	2261	11 ACN89172	Acn89172 Breast ca
38	55	4.6	3113	4 AAS26331	Aas26331 Human cdn
39	55	4.6	3113	8 ABX73672	Abx73672 Human nov
40	54.8	4.6	413	4 AA185315	Aai85315 Human pol
41	54.6	4.5	964	12 ADM47605	Adm47605 Polyucle
42	54.6	4.5	1344	8 ABO77393	Abg77393 Human zin
43	54.4	4.5	344	13 ACN47065	Acn47065 Cotton pr
44	54.4	4.5	858	13 ADR61633	Adr61633 Cotton cd
45	54.4	4.5	1001	3 AAC99900	Aac99900 Human sec

ALIGNMENTS

RESULT 1

AAC50023

ID AAC50023 standard; DNA; 912 BP.

XX AC AAC50023;

XX DT 18-OCT-2000 (first entry)

XX DE Arabidopsis thaliana DNA fragment SEQ ID NO: 63305.

XX KW Hybridisation assay; genetic mapping; gene expression control;

XX KW protein identification; signal transduction pathway; metabolic pathway;

XX KW promoter; termination sequence; ss.

XX OS Arabidopsis thaliana.

XX PN EP1033405-A2.

XX PD 06-SEP-2000.

XX PF 25-FEB-2000; 2000EP-00301439.

XX PR 25-FEB-1999; 99US-0121825P.

XX PR 05-MAR-1999; 99US-0123180P.

XX PR 09-MAR-1999; 99US-0123548P.

XX PR 23-MAR-1999; 99US-0125788P.

XX PR 25-MAR-1999; 99US-0126264P.

XX PR 29-MAR-1999; 99US-0126785P.

XX PR 01-APR-1999; 99US-0127462P.

XX PR 06-APR-1999; 99US-0128234P.

XX PR 08-APR-1999; 99US-0128714P.

XX PR 16-APR-1999; 99US-0129845P.

XX PR 19-APR-1999; 99US-0130077P.

XX PR 21-APR-1999; 99US-0130449P.

XX PR 23-APR-1999; 99US-0130510P.

XX PR 23-APR-1999; 99US-0130891P.

XX PR 28-APR-1999; 99US-0131449P.

XX PR 30-APR-1999; 99US-0132048P.

XX PR 30-APR-1999; 99US-0132407P.

XX PR 04-MAY-1999; 99US-0132484P.

XX PR 05-MAY-1999; 99US-0132485P.

XX PR 06-MAY-1999; 99US-0132486P.

XX PR 06-MAY-1999; 99US-0132487P.

PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 28-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144684P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.

PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.

	22-OCT-1999;	99US-0160980P.		Query Match	20.0%;	Score 239.4;	DB 3;	Length 912;	
	22-OCT-1999;	99US-0160981P.		Best Local Similarity	62.4%;	Pred. No. 1.9e-45;			
	22-OCT-1999;	99US-0160989P.		Matches 375;	Conservative 0;	Mismatches 226;	Indels 0;	Gaps 0;	
PR	25-OCT-1999;	99US-0161404P.							
PR	25-OCT-1999;	99US-0161405P.							
PR	25-OCT-1999;	99US-0161406P.							
PR	26-OCT-1999;	99US-0161359P.							
PR	26-OCT-1999;	99US-0161360P.							
PR	26-OCT-1999;	99US-0161361P.							
PR	28-OCT-1999;	99US-0161920P.							
PR	28-OCT-1999;	99US-0161992P.							
PR	28-OCT-1999;	99US-0161993P.							
PR	29-OCT-1999;	99US-0162142P.							
Qy	382	TGAAGAGAAATCAGAAGATTCTGTTCTACTTTGAACGGGAGGTGTATTTACTCTAGTAG	441						
Db	266	TCAAGAGAAAGCAGAAGAGGTTAAACCGTATTTGAATGGACGACTCTATGTACCTTTGTGCG	325						
Qy	442	GAATGATGGGTTCTGGAAAACTACTGTGGGGAAGATTATGTCTGAAGTCTTGGGTATT	501						
Db	326	GAATGATGGGTTCTGGAAAACAACTGTGGGAAAAGTTAAATGTGCCAAGTGTCTCGGTATA	385						
Qy	502	CGTTCTTTCATAGTCACAAGTTAGTGGAGCAAGCTGTTGGAAATGCCATCAGTTGCCCAA	561						
Db	386	CGTTCTTTGACTGCGACACTTTTGATGAAACAGGCGATGAATGGAACTTCTGTTGCAGAGA	445						
Qy	562	TATTCAAAGTCCATAGTGAAGCCTTCTTTCGGGATAATGAGAGTAGTGTCTTCGAGATT	621						
Db	446	TATTTCTTCATCACGAGAGAAATTTTTTTAGAGGAAAGGAGACCGATGCGCTTTAAGAAGC	505						
Qy	622	TGTCCTCCATGCGACGATTAGTTGTGGCCACCGGAGGTGGTGTGTTATCCGACCAATTA	681						
Db	506	TCTCTTCGAGGTATCAAGTTGTGTTCACAGGTGGAGTGCAGTTATTAAGACCCATT	565						
Qy	682	ACTGGAGATATGAAGAGGGGCCCTATCTGTTTGGTTAGATGTGCCCTTCGATGCTCTTG	741						
Db	566	ACTGGAAATATATGCATAAAGAAATCAGCAATTTGGCTAGATGTGCTCTAGAAGCATTAG	625						
Qy	742	CTAGGCGTATATGCTAAAGTGGAACTGCCTCTCGTCTCTTCTGGACCAACCACTCTGGTG	801						
Db	626	CCCATAGAAATCGCTGCTGTGTGGAACTGATTCACGACCACCTGTACTACGATGAATCAGGAG	685						
Qy	802	ATCCGTACCGAATGGCCCTTTTCTAAGCTCAGCATGCTTGCACAGCAAAAGGGGTGATGCTT	861						
Db	686	ATGCATACTCAGTGGGCTTTCAAACGCTCTCTCGGCTATTTGGGACGACGGCGGTGAAGCAT	745						
Qy	862	ATGCAAAATCAGATGTAAGGGTTTCTCTGGAAAGAGATTGCATGTAAACAAGGTCATGATG	921						
Db	746	ACAAACGCAANTGCCAGTGCTCCTTAGAAATATTTGCAGCAAAAGCGTGGCTATATAA	805						
Qy	922	ATGTCCTTAAGCTGCACCTACTCATGATATTGCAATTGAGTCACTTTCATAAGATCGAGACT	981						
Db	806	ATGTCCTCAGATCTCACCACTGAAATGCTATCGAGGCCCTTCGACCAAGTTCAGCT	865						
Qy	982	T 982							
Db	866	T 866							

RESULT 2	PR	18-JUN-1999;	99US-0139458P.
AAC50026	PR	18-JUN-1999;	99US-0139459P.
ID AAC50026 standard; DNA; 1365 BP.	PR	18-JUN-1999;	99US-0139460P.
XX	PR	18-JUN-1999;	99US-0139461P.
XX	PR	18-JUN-1999;	99US-0139462P.
AC AAC50026;	PR	18-JUN-1999;	99US-0139463P.
XX	PR	18-JUN-1999;	99US-0139750P.
DT 18-OCT-2000 (first entry)	PR	18-JUN-1999;	99US-0139763P.
XX	PR	21-JUN-1999;	99US-0139817P.
XX	PR	22-JUN-1999;	99US-0139899P.
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 63316.	PR	23-JUN-1999;	99US-0140353P.

PR	23-JUN-1999;	99US-0140354P.	PR	01-SEP-1999;	99US-0151930P.
PR	24-JUN-1999;	99US-0140695P.	PR	07-SEP-1999;	99US-0152363P.
PR	28-JUN-1999;	99US-0140823P.	PR	10-SEP-1999;	99US-0153070P.
PR	29-JUN-1999;	99US-0140991P.	PR	13-SEP-1999;	99US-0153758P.
PR	30-JUN-1999;	99US-0141287P.	PR	15-SEP-1999;	99US-0154018P.
PR	01-JUL-1999;	99US-0141842P.	PR	16-SEP-1999;	99US-0154039P.
PR	01-JUL-1999;	99US-0142154P.	PR	20-SEP-1999;	99US-0154779P.
PR	02-JUL-1999;	99US-0142055P.	PR	22-SEP-1999;	99US-0155113P.
PR	06-JUL-1999;	99US-0142390P.	PR	23-SEP-1999;	99US-0155486P.
PR	08-JUL-1999;	99US-0142803P.	PR	24-SEP-1999;	99US-0155659P.
PR	09-JUL-1999;	99US-0142920P.	PR	28-SEP-1999;	99US-0156458P.
PR	12-JUL-1999;	99US-0142977P.	PR	29-SEP-1999;	99US-0156596P.
PR	13-JUL-1999;	99US-0143542P.	PR	04-OCT-1999;	99US-0157117P.
PR	14-JUL-1999;	99US-0143624P.	PR	05-OCT-1999;	99US-0157753P.
PR	15-JUL-1999;	99US-0144005P.	PR	06-OCT-1999;	99US-0157865P.
PR	16-JUL-1999;	99US-0144086P.	PR	07-OCT-1999;	99US-0158029P.
PR	16-JUL-1999;	99US-0144086P.	PR	08-OCT-1999;	99US-0158232P.
PR	19-JUL-1999;	99US-0144325P.	PR	12-OCT-1999;	99US-0158369P.
PR	19-JUL-1999;	99US-0144331P.	PR	13-OCT-1999;	99US-0159293P.
PR	19-JUL-1999;	99US-0144332P.	PR	13-OCT-1999;	99US-0159294P.
PR	19-JUL-1999;	99US-0144333P.	PR	13-OCT-1999;	99US-0159295P.
PR	19-JUL-1999;	99US-0144334P.	PR	14-OCT-1999;	99US-0159329P.
PR	19-JUL-1999;	99US-0144335P.	PR	14-OCT-1999;	99US-0159330P.
PR	20-JUL-1999;	99US-0144352P.	PR	14-OCT-1999;	99US-0159331P.
PR	20-JUL-1999;	99US-0144632P.	PR	14-OCT-1999;	99US-0159637P.
PR	20-JUL-1999;	99US-0144884P.	PR	14-OCT-1999;	99US-0159638P.
PR	21-JUL-1999;	99US-0144814P.	PR	18-OCT-1999;	99US-0159584P.
PR	21-JUL-1999;	99US-0145086P.	PR	21-OCT-1999;	99US-0160741P.
PR	21-JUL-1999;	99US-0145088P.	PR	21-OCT-1999;	99US-0160767P.
PR	21-JUL-1999;	99US-0145088P.	PR	21-OCT-1999;	99US-0160768P.
PR	22-JUL-1999;	99US-0145087P.	PR	21-OCT-1999;	99US-0160770P.
PR	22-JUL-1999;	99US-0145089P.	PR	21-OCT-1999;	99US-0160814P.
PR	22-JUL-1999;	99US-0145192P.	PR	21-OCT-1999;	99US-0160815P.
PR	23-JUL-1999;	99US-0145145P.	PR	22-OCT-1999;	99US-0160980P.
PR	23-JUL-1999;	99US-0145218P.	PR	22-OCT-1999;	99US-0160981P.
PR	23-JUL-1999;	99US-0145224P.	PR	22-OCT-1999;	99US-0160989P.
PR	26-JUL-1999;	99US-0145276P.	PR	25-OCT-1999;	99US-0161404P.
PR	27-JUL-1999;	99US-0145913P.	PR	25-OCT-1999;	99US-0161405P.
PR	27-JUL-1999;	99US-0145918P.	PR	25-OCT-1999;	99US-0161406P.
PR	27-JUL-1999;	99US-0145919P.	PR	26-OCT-1999;	99US-0161359P.
PR	28-JUL-1999;	99US-0145951P.	PR	26-OCT-1999;	99US-0161360P.
PR	02-AUG-1999;	99US-0146386P.	PR	26-OCT-1999;	99US-0161361P.
PR	02-AUG-1999;	99US-0146388P.	PR	28-OCT-1999;	99US-0161920P.
PR	03-AUG-1999;	99US-0147038P.	PR	28-OCT-1999;	99US-0161992P.
PR	04-AUG-1999;	99US-0147204P.	PR	28-OCT-1999;	99US-0161993P.
PR	04-AUG-1999;	99US-0147302P.	PR	29-OCT-1999;	99US-0162142P.
PR	05-AUG-1999;	99US-0147192P.	Query Match		
PR	05-AUG-1999;	99US-0147260P.	Best Local Similarity		
PR	06-AUG-1999;	99US-0147303P.	62.2%; Pred. No. 8.7e-45;		
PR	06-AUG-1999;	99US-0147416P.	Matches 373; Conservative		
PR	09-AUG-1999;	99US-0147493P.	0; Mismatches 227; Indels		
PR	09-AUG-1999;	99US-0147935P.	0; Gaps		
PR	10-AUG-1999;	99US-0148171P.	0;		
PR	11-AUG-1999;	99US-0148319P.	QY	383	GAAGAGAAATCAGAAGAAGTTCTGTTCTACTTGAACGGAGGTTATTTACCTAGTAGG 442
PR	12-AUG-1999;	99US-0148341P.	Db	377	GCAAGAAGCAGAGAAGGTTAAACCGTATTTTGAATGGAACGATCTATGTACCTTGTCCG 436
PR	13-AUG-1999;	99US-0148565P.	QY	443	AATGATGGTTCTCGAAAAAGTACTGTGGGAAGATTATGTCTGAAGTCTTCGGTTATTC 502
PR	13-AUG-1999;	99US-0148684P.	Db	437	AATGATGGTTCTCGGAAAAACACTGTGGGAAAGTTAATGTCCAAGTGTCTGGTTATAC 496
PR	16-AUG-1999;	99US-0149368P.	QY	503	GTTCTTTGATAGTGAAGTTAGTGGAGCAAGCTCTTTGGAAATGCCATCAGTTGCCCAAT 562
PR	17-AUG-1999;	99US-0149175P.	Db	497	GTTCTTTGACTCGGACACTTTTGATTGAACAGGCGATGAATGGAATCTCTGTTGCAGAGAT 556
PR	20-AUG-1999;	99US-0149723P.	QY	563	ATTCAAGGTCCATAGTAGAGCCCTTTTCGGGATATAGAGAGTACTGTCTGAGAGATTT 622
PR	20-AUG-1999;	99US-0149929P.	Db	557	ATTTGTTTCATCAGGAGAGAAATTTTTTAGAGNAAAGGAGACCGATGCGCTTTAAGAAGCT 616
PR	23-AUG-1999;	99US-0149930P.	QY	623	GTCCTCCATCGACGATTAGTTGTCACCGAGGTGGTGTCTTATCCGACCAATTAA 682
PR	25-AUG-1999;	99US-0150566P.	Db	617	CTCTTCGAGGTATCAAGTTGTTGTTTCCACAGTGGAGGTGCAGTTATAAGACCCATTAA 676
PR	26-AUG-1999;	99US-0150884P.	QY	683	CTTGAGATATATGAAGAGGGGCTATCTGTTTGGTTAGATGTCCTTGGATGCTCTTCG 742
PR	27-AUG-1999;	99US-0151065P.			
PR	27-AUG-1999;	99US-0151066P.			
PR	27-AUG-1999;	99US-0151080P.			
PR	30-AUG-1999;	99US-0151303P.			
PR	31-AUG-1999;	99US-0151438P.			

PR	03-AUG-1999;	99US-0147038P.	PR	28-OCT-1999;	99US-0161993P.	Query Match	19.7%;	Score	236.8;	DB	3;	Length	1491;	
PR	04-AUG-1999;	99US-0147204P.	PR	29-OCT-1999;	99US-0162142P.	Best Local Similarity	62.2%;	Pred. No.	8.9e-45;	Mismatches	227;	Indels	0;	
PR	04-AUG-1999;	99US-0147302P.				Matches	373;	Conservative	0;				Gaps	0;
PR	05-AUG-1999;	99US-0147192P.												
PR	05-AUG-1999;	99US-0147260P.												
PR	06-AUG-1999;	99US-0147303P.												
PR	06-AUG-1999;	99US-0147416P.												
PR	09-AUG-1999;	99US-0147493P.												
PR	09-AUG-1999;	99US-0147935P.												
PR	10-AUG-1999;	99US-0148171P.												
PR	11-AUG-1999;	99US-0148319P.												
PR	12-AUG-1999;	99US-0148341P.												
PR	13-AUG-1999;	99US-0148565P.												
PR	13-AUG-1999;	99US-0148684P.												
PR	16-AUG-1999;	99US-0149368P.												
PR	17-AUG-1999;	99US-0149175P.												
PR	18-AUG-1999;	99US-0149426P.												
PR	20-AUG-1999;	99US-0149722P.												
PR	20-AUG-1999;	99US-0149723P.												
PR	20-AUG-1999;	99US-0149929P.												
PR	23-AUG-1999;	99US-0149902P.												
PR	23-AUG-1999;	99US-0149930P.												
PR	25-AUG-1999;	99US-0150566P.												
PR	26-AUG-1999;	99US-0150884P.												
PR	27-AUG-1999;	99US-0151066P.												
PR	27-AUG-1999;	99US-0151080P.												
PR	30-AUG-1999;	99US-0151303P.												
PR	31-AUG-1999;	99US-0151438P.												
PR	01-SEP-1999;	99US-0151930P.												
PR	07-SEP-1999;	99US-0152363P.												
PR	10-SEP-1999;	99US-0153070P.												
PR	13-SEP-1999;	99US-0153758P.												
PR	15-SEP-1999;	99US-0154018P.												
PR	16-SEP-1999;	99US-0154039P.												
PR	20-SEP-1999;	99US-0154779P.												
PR	22-SEP-1999;	99US-0155139P.												
PR	23-SEP-1999;	99US-0155486P.												
PR	24-SEP-1999;	99US-0155659P.												
PR	28-SEP-1999;	99US-0156458P.												
PR	29-SEP-1999;	99US-0156596P.												
PR	04-OCT-1999;	99US-0157117P.												
PR	05-OCT-1999;	99US-0157753P.												
PR	06-OCT-1999;	99US-0157865P.												
PR	07-OCT-1999;	99US-0158029P.												
PR	08-OCT-1999;	99US-0158232P.												
PR	12-OCT-1999;	99US-0158369P.												
PR	13-OCT-1999;	99US-0159293P.												
PR	13-OCT-1999;	99US-0159294P.												
PR	13-OCT-1999;	99US-0159295P.												
PR	14-OCT-1999;	99US-0159329P.												
PR	14-OCT-1999;	99US-0159330P.												
PR	14-OCT-1999;	99US-0159331P.												
PR	14-OCT-1999;	99US-0159637P.												
PR	14-OCT-1999;	99US-0159638P.												
PR	18-OCT-1999;	99US-0159584P.												
PR	21-OCT-1999;	99US-0160741P.												
PR	21-OCT-1999;	99US-0160767P.												
PR	21-OCT-1999;	99US-0160768P.												
PR	21-OCT-1999;	99US-0160770P.												
PR	21-OCT-1999;	99US-0160814P.												
PR	21-OCT-1999;	99US-0160815P.												
PR	22-OCT-1999;	99US-0160980P.												
PR	22-OCT-1999;	99US-0160981P.												
PR	22-OCT-1999;	99US-0160989P.												
PR	25-OCT-1999;	99US-0161404P.												
PR	25-OCT-1999;	99US-0161405P.												
PR	25-OCT-1999;	99US-0161406P.												
PR	26-OCT-1999;	99US-0161359P.												
PR	26-OCT-1999;	99US-0161360P.												
PR	26-OCT-1999;	99US-0161361P.												
PR	28-OCT-1999;	99US-0161920P.												
PR	28-OCT-1999;	99US-0161922P.												
PR	03-AUG-1999;	99US-0147038P.	PR	28-OCT-1999;	99US-0161993P.	Query Match	19.7%;	Score	236.8;	DB	3;	Length	1491;	
PR	04-AUG-1999;	99US-0147204P.	PR	29-OCT-1999;	99US-0162142P.	Best Local Similarity	62.2%;	Pred. No.	8.9e-45;	Mismatches	227;	Indels	0;	
PR	04-AUG-1999;	99US-0147302P.				Matches	373;	Conservative	0;				Gaps	0;
PR	05-AUG-1999;	99US-0147192P.												
PR	05-AUG-1999;	99US-0147260P.												
PR	06-AUG-1999;	99US-0147303P.												
PR	06-AUG-1999;	99US-0147416P.												
PR	09-AUG-1999;	99US-0147493P.												
PR	09-AUG-1999;	99US-0147935P.												
PR	10-AUG-1999;	99US-0148171P.												
PR	11-AUG-1999;	99US-0148319P.												
PR	12-AUG-1999;	99US-0148341P.												
PR	13-AUG-1999;	99US-0148565P.												
PR	13-AUG-1999;	99US-0148684P.												
PR	16-AUG-1999;	99US-0149368P.												
PR	17-AUG-1999;	99US-0149175P.												
PR	18-AUG-1999;	99US-0149426P.												
PR	20-AUG-1999;	99US-0149722P.												
PR	20-AUG-1999;	99US-0149723P.												
PR	20-AUG-1999;	99US-0149929P.												
PR	23-AUG-1999;	99US-0149902P.												
PR	23-AUG-1999;	99US-0149930P.												
PR	25-AUG-1999;	99US-0150566P.												
PR	26-AUG-1999;	99US-0150884P.												
PR	27-AUG-1999;	99US-0151066P.												
PR	27-AUG-1999;	99US-0151080P.												
PR	30-AUG-1999;	99US-0151303P.												
PR	31-AUG-1999;	99US-0151438P.												
PR	01-SEP-1999;	99US-0151930P.												
PR	07-SEP-1999;	99US-0152363P.												
PR	10-SEP-1999;	99US-0153070P.												
PR	13-SEP-1999;	99US-0153758P.												
PR	15-SEP-1999;	99US-0154018P.												
PR	16-SEP-1999;	99US-0154039P.												
PR	20-SEP-1999;	99US-0154779P.												
PR	22-SEP-1999;	99US-0155139P.												
PR	23-SEP-1999;	99US-0155486P.												
PR	24-SEP-1999;	99US-0155659P.												
PR	28-SEP-1999;	99US-0156458P.												
PR	29-SEP-1999;	99US-0156596P.												
PR	04-OCT-1999;	99US-0157117P.												
PR	05-OCT-1999;	99US-0157753P.												
PR	06-OCT-1999;	99US-0157865P.												
PR	07-OCT-1999;	99US-0158029P.												
PR	08-OCT-1999;	99US-0158232P.												
PR	12-OCT-1999;	99US-0158369P.												
PR	13-OCT-1999;	99US-0159293P.												
PR	13-OCT-1999;	99US-0159294P.												
PR	13-OCT-1999;	99US-0159295P.												
PR	14-OCT-1999;	99US-0159329P.												
PR	14-OCT-1999;	99US-0159330P.												
PR	14-OCT-1999;	99US-0159331P.												
PR	14-OCT-1999;	99US-0159637P.												
PR	14-OCT-1999;	99US-0159638P.												
PR	18-OCT-1999;	99US-0159584P.												
PR	21-OCT-1999;	99US-0160741P.												
PR	21-OCT-1999;	99US-0160767P.												
PR	21-OCT-1999;	99US-0160768P.												
PR	21-OCT-1999;	99US-0160770P.												
PR	21-OCT-1999;	99US-0160814P.												
PR	21-OCT-1999;	99US-0160815P.												
PR	22-OCT-1999;	99US-0160980P.												
PR	22-OCT-1999;	99US-0160981P.												

PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131149P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-01334256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134270P.
PR 18-MAY-1999; 99US-0134370P.
PR 19-MAY-1999; 99US-0134768P.
PR 20-MAY-1999; 99US-0134941P.
PR 21-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 21-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 24-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.

PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145911P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 05-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156450P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.

PR	07-OCT-1999;	99US-0158029P.	
PR	08-OCT-1999;	99US-0158232P.	
PR	12-OCT-1999;	99US-0158369P.	
PR	13-OCT-1999;	99US-0159293P.	
PR	13-OCT-1999;	99US-0159294P.	
PR	13-OCT-1999;	99US-0159295P.	
PR	14-OCT-1999;	99US-0159329P.	
PR	14-OCT-1999;	99US-0159330P.	
PR	14-OCT-1999;	99US-0159331P.	
PR	14-OCT-1999;	99US-0159637P.	
PR	14-OCT-1999;	99US-0159638P.	
PR	18-OCT-1999;	99US-0159584P.	
PR	21-OCT-1999;	99US-0160741P.	
PR	21-OCT-1999;	99US-0160767P.	
PR	21-OCT-1999;	99US-0160768P.	
PR	21-OCT-1999;	99US-0160770P.	
PR	21-OCT-1999;	99US-0160814P.	
PR	21-OCT-1999;	99US-0160815P.	
PR	22-OCT-1999;	99US-0160980P.	
PR	22-OCT-1999;	99US-0160981P.	
PR	22-OCT-1999;	99US-0160989P.	
PR	25-OCT-1999;	99US-0161404P.	
PR	25-OCT-1999;	99US-0161405P.	
PR	25-OCT-1999;	99US-0161406P.	
PR	26-OCT-1999;	99US-0161359P.	
PR	26-OCT-1999;	99US-0161360P.	
PR	26-OCT-1999;	99US-0161361P.	
PR	28-OCT-1999;	99US-0161920P.	
PR	28-OCT-1999;	99US-0161992P.	
PR	28-OCT-1999;	99US-0161993P.	
PR	29-OCT-1999;	99US-0162142P.	
Query Match 19.7%; Score 236.2; DB 3; Length 1363;			
Best Local Similarity 62.1%; Pred. No. 1.2e-44;			
Matches 373; Conservative 0; Mismatches 228; Indels 0; Gaps 0;			
QY	382	TGAAGAGAAATCAGAAAGTTCTCTTCTACTTGAACGGAGGTGATTATACCTAGTAG	441
DB	372	TCAGAGAAAGCAGAAAGAGTTAAACCGTATTTGAATGACGATCTATGTAACCTTGTGC	431
QY	442	GAATGATGGTTCTGGAAAAAGTACTGTGGGGAAGATTATGTCGAACTCTTGGGTATT	501
DB	432	GAATGATGGTTCTGGAAAAACAACCTGTGGGAAAGTTAATGTCCAAAGTCTCGGTTATA	491
QY	502	CGTCTTTCATAGTACRAGTTAGTGGACGACGCTGTTGGAATGCCATCAGTTGCCRAA	561
DB	492	CGTCTTTCATAGTACRAGTTAGTGGACGACGCTGTTGGAATGCCATCAGTTGCCRAA	551
QY	562	TATTCAGAGTCCATAGTGAAGCCTTCTTTCGGGATATAGAGATAGTGTCTTGAGAGATT	621
DB	552	TATTCAGTTCATCAGGAGAGATTTTTCAGAGAAAGGAGACCGATCGCTTAAGAAGC	611
QY	622	TGTCCTCCATGCGACGATAGTTGTTGCCACCGAGGTGGTGTCTGTTATCCGACCAATTA	681
DB	612	TCTCTTCGAGGTATCAAGTTGTTTCCACAGGTGGAGTGCAGTGTATAAGACCCATTA	671
QY	682	ACTGGAGATATAGAGAGGGGCTATCTGTTGGTATAGTGTGCCCTTGGATGCTCTTG	741
DB	672	ACTGGAAGTATATGTCATAAGGAATCAGCATTTTGGCTAGATGTGCTCTAGAGCAATTAG	731
QY	742	CTAGGCGTATTGTAAGTGGGAACCTCTCGTCTCTTCTGGACCAACCATCTGCTG	801
DB	732	CCCATAGATCGTGTCTGTTGGAACTGATTCAGACCACTGCTACACGATGATCAGGAG	791
QY	802	ATCCGTACGGAATGGCCTTTTCTTAAGCTCAGCATGCTTGACAGCAAAAGGGGTGATG	861
DB	792	ATGCATACCTCAGTGGCTTTTCAAAACGCTCTCTCGGCTATTTGGGACGACGCGGTGAAGCAT	851
QY	862	ATCCAAATCAGATTAAGGTTTCTCTGGAAGATTCGATGTAACAAAGTCTCATGATG	921
DB	852	ACAAACCGCAATGCCAGATCTCTCTTAGAAATATTCAGCAAAAGCGTGGCTATATAA	911
QY	922	ATGTCCTTAAGCTGACACCTACTCTATATATGCAATTCAGTCACTTTCATAAGATCGAGCT	981

DB	912	ATGTCCTCAGATCTCACCAACTGAATTTGTATGAGCCTTCTGAGCAAGTCTTGAGCT	971
QY	982	T 982	
DB	972	T 972	
RESULT 5			
ID	ADKS2674	standard; DNA; 637 BP.	
XX	ADKS2674;		
XX	06-MAY-2004	(first entry)	
XX	Plant DNA sequence which confers altered metabolic characteristic #57.		
XX	altered metabolic characteristic; plant; acid metabolism;		
KW	alcohol metabolism; fatty acid metabolism;		
KW	branched fatty acid metabolism; alkaloid metabolism;		
KW	amino acid metabolism; ester metabolism; glyceride metabolism;		
KW	phenolic metabolism; carbohydrate metabolism; steroid metabolism;		
KW	terpene metabolism; isoprenoid metabolism; alkene metabolism;		
KW	alkyne metabolism; hydrocarbon metabolism; ketone metabolism;		
KW	quinone metabolism; disease resistance; gene shuffling; sexual PCR; ds.		
OS	Nicotiana benthamiana.		
XX	WO2003020936-A1.		
XX	13-MAR-2003.		
XX	30-AUG-2002; 2002WO-US027884.		
XX	31-AUG-2001; 2001US-0316471P.		
PA	(DOWC) DOW CHEM CO.		
PA	(DOWC) DOW AGROSCIENCES LLC.		
XX	Weglarz T, Gachotte D, Blakeslee B, McCreary DA, Pell RJ;		
PI	Oriado JVB, Crosley R, Reddy AS, Shukla V, Larrinua I, Miller BA;		
XX	WPI; 2003-313091/30.		
DR	Novel genes that confer altered metabolic characteristics in Nicotiana		
PT	benthamiana plants, useful for altering the levels of metabolites e.g.		
PT	acids, fatty acids, amino acids, carbohydrates, hydrocarbons and sterols.		
XX	Claim 1; SEQ ID NO 57; 2576pp; English.		
XX	The invention comprises DNA sequences which confer an altered metabolic		
CC	characteristic when they are expressed in a plant. The DNA sequences of		
CC	the invention are useful for producing plants with an altered metabolic		
CC	characteristic, such as: altered acid metabolism, alcohol metabolism,		
CC	fatty acid metabolism, branched fatty acid metabolism, alkaloid or other		
CC	base metabolism, altered amino acid metabolism, altered ester metabolism,		
CC	altered glyceride metabolism, altered phenolic metabolism, altered		
CC	carbohydrate metabolism, altered steroid, oxygenated terpene, or		
CC	isoprenoid metabolism, alkene or alkyne metabolism, hydrocarbon		
CC	metabolism, ketone or quinone metabolism. The DNA sequences of the		
CC	invention may be used to provide disease resistance in a plant and gene		
CC	shuffling or sexual PCR procedures. The present nucleic acid represents a		
CC	DNA sequence of the invention.		
XX	Sequence 637 BP; 185 A; 103 C; 165 G; 184 T; 0 U; 0 Other;		
Query Match 19.5%; Score 233.6; DB 10; Length 637;			
Best Local Similarity 63.7%; Pred. No. 3.8e-44;			
Matches 372; Conservative 0; Mismatches 209; Indels 3; Gaps 1;			
QY	411	TACTTGAACGGAGGTGTTATTTACCTAGTAGGAATGATGGTCTCGAAAAAGTACTGTG	470

Db	9	TATCTAGATGGACGATGTATATACCTCGTTGGAGTGATGGCTCTGCGAAAAACAAC	68	PR	23-APR-1999;	99US-0130510P.
Qy	471	GGGAAGATTATGTCTGAAGTCTTTGGGTTATTCGTTCTTTTATAGTACACAGTTAGTGGAG	530	PR	23-APR-1999;	99US-0130891P.
Db	69	GGCGTATTTTGGCAGAAACACTGGGATATTCCTTTTGGACTGTGACAGGCTGATAGAG	128	PR	28-APR-1999;	99US-0131449P.
Qy	531	CAAGCTGTTGGATGCCATCAGTTGGCCAAATATTCGAAGTCCATAGTGAAGCCTTCTTTT	590	PR	30-APR-1999;	99US-0132048P.
Db	129	CAGGCTGTTGGTGGAACTACAGTAGCTGAAATCTTCAAGCTTCGTGGAGAGAGCTTCTTT	188	PR	04-MAY-1999;	99US-0132407P.
Qy	591	CGGGAATAACAGAGTAGTGTCTTGAGAGATTTGTCTCCATCGAGCATTTAGTTGTTGCC	650	PR	05-MAY-1999;	99US-0132484P.
Db	189	AGGACAAATGAGACGGAGGTATTGCACAGCTGTCTGCGATGCATCGCTTGTGTTCAT	248	PR	06-MAY-1999;	99US-0132485P.
Qy	651	ACCGGAGGTGCTGCTTTATCCGACCAATTAACCTGGAGATATATGAAGAGGGGCCCTATCT	710	PR	06-MAY-1999;	99US-0132486P.
Db	249	ACAGGTGGAGTGCAGTTGTTCTGCCATTTAATTGGAGACATATGCACAGGCTATTAGT	308	PR	07-MAY-1999;	99US-0132863P.
Qy	711	GTTTGGTTAGATGGCCTTGGATGCTCTTGGCTAGGGGTATTCGTTAAAGTGGGAACCTGCC	770	PR	11-MAY-1999;	99US-0132864P.
Db	309	GTTTGGTTAGATGTTCTTTTAGATGCTTTGGCCCAAGAGAGTTACTGCTGAAGGAACATAA	368	PR	14-MAY-1999;	99US-0134218P.
Qy	771	TCTCGTCTCTCTTGGACCAACCATCTGGTGATCCGTACGCAATGCGCTTTTCTAAGCTC	830	PR	14-MAY-1999;	99US-0134219P.
Db	369	TCTCGACCCCTATTACATGAAGAATCAGGAGACATTTATGATAAGACTTTGAAGCGGTTA	428	PR	14-MAY-1999;	99US-0134221P.
Qy	831	AGCATGCTTGCACAGCAAGGGGTGATGCTTTATGCAAAATGCAGATGTAAAGGGTTTCTCTG	890	PR	14-MAY-1999;	99US-0134221P.
Db	429	ACTACTTTAATGAGACAGGGGTGAAACTATGCCAATGCAAGTGTACGGGTTTCACTA	488	PR	18-MAY-1999;	99US-0134768P.
Qy	891	GAAGAGATTGCATGTAAACAAAGTCTATGATGATGCTCTTAAGCTGCACCTACTGATATT	950	PR	19-MAY-1999;	99US-0134768P.
Db	489	GAAATATTTCAGTGAAGAGGGG---AAAAGATGCTCTGCCATATTACACCTACTGAAATT	545	PR	20-MAY-1999;	99US-0135124P.
Qy	951	GCAATTGAGTCATTCATAGATCGAGAGCTTCGTATCGAGCA	994	PR	21-MAY-1999;	99US-0135353P.
Db	546	ACTCTAGAGTTCTTATACAAATTCAGAACTTCTTAAAGAAACA	589	PR	24-MAY-1999;	99US-0135629P.
RESULT 6						
AAC40184						
XX	AAC40184;					
XX	17-OCT-2000 (first entry)					
XX	Arabidopsis thaliana DNA fragment SEQ ID NO: 27368.					
XX	Hybridisation assay; genetic mapping; gene expression control;					
KW	protein identification; signal transduction pathway; metabolic pathway;					
KW	promoter; termination sequence; ss.					
XX	Arabidopsis thaliana.					
XX	EP1033405-A2.					
XX	06-SEP-2000.					
XX	25-FEB-2000; 2000EP-00301439.					
XX	25-FEB-1999; 99US-0121825P.					
PR	05-MAR-1999; 99US-0123180P.					
PR	09-MAR-1999; 99US-0123548P.					
PR	23-MAR-1999; 99US-0125788P.					
PR	25-MAR-1999; 99US-0126264P.					
PR	29-MAR-1999; 99US-0126785P.					
PR	01-APR-1999; 99US-0127462P.					
PR	06-APR-1999; 99US-0128234P.					
PR	08-APR-1999; 99US-0128714P.					
PR	16-APR-1999; 99US-0129845P.					
PR	19-APR-1999; 99US-0130077P.					
PR	21-APR-1999; 99US-0130449P.					

PR	20-JUL-1999;	99US-0144632P.	PR	14-OCT-1999;	99US-0159637P.
PR	20-JUL-1999;	99US-0144884P.	PR	14-OCT-1999;	99US-0159638P.
PR	21-JUL-1999;	99US-0144814P.	PR	18-OCT-1999;	99US-0159584P.
PR	21-JUL-1999;	99US-0145086P.	PR	21-OCT-1999;	99US-0160741P.
PR	21-JUL-1999;	99US-0145088P.	PR	21-OCT-1999;	99US-0160767P.
PR	22-JUL-1999;	99US-0145085P.	PR	21-OCT-1999;	99US-0160768P.
PR	22-JUL-1999;	99US-0145087P.	PR	21-OCT-1999;	99US-0160770P.
PR	22-JUL-1999;	99US-0145089P.	PR	21-OCT-1999;	99US-0160814P.
PR	22-JUL-1999;	99US-0145192P.	PR	21-OCT-1999;	99US-0160815P.
PR	23-JUL-1999;	99US-0145145P.	PR	22-OCT-1999;	99US-0160980P.
PR	23-JUL-1999;	99US-0145218P.	PR	22-OCT-1999;	99US-0160981P.
PR	23-JUL-1999;	99US-0145224P.	PR	22-OCT-1999;	99US-0160989P.
PR	26-JUL-1999;	99US-0145276P.	PR	25-OCT-1999;	99US-0161404P.
PR	27-JUL-1999;	99US-0145913P.	PR	25-OCT-1999;	99US-0161405P.
PR	27-JUL-1999;	99US-0145918P.	PR	25-OCT-1999;	99US-0161406P.
PR	27-JUL-1999;	99US-0145919P.	PR	26-OCT-1999;	99US-0161359P.
PR	28-JUL-1999;	99US-0145951P.	PR	26-OCT-1999;	99US-0161360P.
PR	02-AUG-1999;	99US-0146386P.	PR	26-OCT-1999;	99US-0161361P.
PR	02-AUG-1999;	99US-0146388P.	PR	28-OCT-1999;	99US-0161920P.
PR	02-AUG-1999;	99US-0146389P.	PR	28-OCT-1999;	99US-0161992P.
PR	03-AUG-1999;	99US-0147038P.	PR	28-OCT-1999;	99US-0161993P.
PR	04-AUG-1999;	99US-0147204P.	PR	29-OCT-1999;	99US-0162142P.
PR	04-AUG-1999;	99US-0147302P.	Query Match 18.9%; Score 227; DB 3; Length 1484;		
PR	05-AUG-1999;	99US-0147192P.	Best Local Similarity 62.2%; Pred. No. 1.7e-42;		
PR	05-AUG-1999;	99US-0147260P.	Matches 374; Conservative 0; Mismatches 225; Indels 2; Gaps 1;		
PR	06-AUG-1999;	99US-0147303P.			
PR	06-AUG-1999;	99US-0147416P.			
PR	09-AUG-1999;	99US-0147493P.			
PR	09-AUG-1999;	99US-0147935P.			
PR	10-AUG-1999;	99US-0148171P.			
PR	11-AUG-1999;	99US-0148319P.			
PR	12-AUG-1999;	99US-0148341P.			
PR	13-AUG-1999;	99US-0148565P.			
PR	13-AUG-1999;	99US-0148684P.			
PR	16-AUG-1999;	99US-0149368P.			
PR	17-AUG-1999;	99US-0149175P.			
PR	18-AUG-1999;	99US-0149426P.			
PR	20-AUG-1999;	99US-0149722P.			
PR	20-AUG-1999;	99US-0149723P.			
PR	20-AUG-1999;	99US-0149929P.			
PR	23-AUG-1999;	99US-0149902P.			
PR	23-AUG-1999;	99US-0149930P.			
PR	25-AUG-1999;	99US-0150566P.			
PR	26-AUG-1999;	99US-0150884P.			
PR	27-AUG-1999;	99US-0151065P.			
PR	27-AUG-1999;	99US-0151066P.			
PR	30-AUG-1999;	99US-0151080P.			
PR	30-AUG-1999;	99US-0151303P.			
PR	31-AUG-1999;	99US-0151438P.			
PR	01-SEP-1999;	99US-0151930P.			
PR	07-SEP-1999;	99US-0152363P.			
PR	10-SEP-1999;	99US-0153070P.			
PR	13-SEP-1999;	99US-0153758P.			
PR	15-SEP-1999;	99US-0154018P.			
PR	16-SEP-1999;	99US-0154039P.			
PR	20-SEP-1999;	99US-0154779P.			
PR	22-SEP-1999;	99US-0155139P.			
PR	23-SEP-1999;	99US-0155486P.			
PR	24-SEP-1999;	99US-0155659P.			
PR	28-SEP-1999;	99US-0156458P.			
PR	29-SEP-1999;	99US-0156596P.			
PR	04-OCT-1999;	99US-0157117P.			
PR	05-OCT-1999;	99US-0157753P.			
PR	06-OCT-1999;	99US-0157865P.			
PR	07-OCT-1999;	99US-0158029P.			
PR	08-OCT-1999;	99US-0158232P.			
PR	12-OCT-1999;	99US-0158369P.			
PR	13-OCT-1999;	99US-0159293P.			
PR	13-OCT-1999;	99US-0159294P.			
PR	13-OCT-1999;	99US-0159295P.			
PR	14-OCT-1999;	99US-0159329P.			
PR	14-OCT-1999;	99US-0159330P.			
PR	14-OCT-1999;	99US-0159331P.			
QY	382	TGAAGAGAAATCAGAAGATTCTGTTCTACTTTGAACGGGAGGTGTATTACCTAGTAG	441	382	TGAAGAGAAATCAGAAGATTCTGTTCTACTTTGAACGGGAGGTGTATTACCTAGTAG
DB	543	TCACAGAGAAAAGCAGAGAGGTTAAACCGTATTTCGAATGGACGATCTATGTACCTTTGCG	602	543	TCACAGAGAAAAGCAGAGAGGTTAAACCGTATTTCGAATGGACGATCTATGTACCTTTGCG
QY	442	GAATGATCGGTTCTTGAAAAAGTACTGTGGGGAAGATTATGCTCAAGTCTTGGGTATT	501	442	GAATGATCGGTTCTTGAAAAAGTACTGTGGGGAAGATTATGCTCAAGTCTTGGGTATT
DB	603	GAATGATCGGTTCTTGGGAAAAACAACGTGGGAAAAGTTAATGTCCAAAGTGTCTCGTTATA	662	603	GAATGATCGGTTCTTGGGAAAAACAACGTGGGAAAAGTTAATGTCCAAAGTGTCTCGTTATA
QY	502	CGTTCTTTGATAGTGACAAAGTTAGTGGAGCAAGCTGTTGGAAATGCCATCAGTTGCCAAA	561	502	CGTTCTTTGATAGTGACAAAGTTAGTGGAGCAAGCTGTTGGAAATGCCATCAGTTGCCAAA
DB	663	CGTTCTTTGACTGCGACACTTTGATTGAGGGCGATGAATGGA--ACTTCTGTTGCAGAGA	720	663	CGTTCTTTGACTGCGACACTTTGATTGAGGGCGATGAATGGA--ACTTCTGTTGCAGAGA
QY	562	TATTCAGGTCCATAGTGAAGCTTCTTTCGGGATAATGAGAGTAGTGTCTTGAGAGATT	621	562	TATTCAGGTCCATAGTGAAGCTTCTTTCGGGATAATGAGAGTAGTGTCTTGAGAGATT
DB	721	TATTTGTTTCATCAGCGAGAGAAATTTTTTAGAGGAAAGGAGACCGATGCGCTTAAGAAGC	780	721	TATTTGTTTCATCAGCGAGAGAAATTTTTTAGAGGAAAGGAGACCGATGCGCTTAAGAAGC
QY	622	TGTCCTCCATGCGACGATTAGTTGTTGCCACCGGAGGTGCTGCTGTATCCGACCAATTA	681	622	TGTCCTCCATGCGACGATTAGTTGTTGCCACCGGAGGTGCTGCTGTATCCGACCAATTA
DB	781	TCTCTTCGAGGTATCAAGTTGTTGTTTCCACAGGTGGAGGTGCAGTTATAAGACCATTA	840	781	TCTCTTCGAGGTATCAAGTTGTTGTTTCCACAGGTGGAGGTGCAGTTATAAGACCATTA
QY	682	ACTGGAGATATATGAAGAGGGGCCCTATCTGTTTGGTTAGATGTGCCCTTGGATGCTCTTG	741	682	ACTGGAGATATATGAAGAGGGGCCCTATCTGTTTGGTTAGATGTGCCCTTGGATGCTCTTG
DB	841	ACTGGAAGTATATGCATAAAGGAATCAGCAATTTGGTAGATGTGCTCTAGAAGCATTAG	900	841	ACTGGAAGTATATGCATAAAGGAATCAGCAATTTGGTAGATGTGCTCTAGAAGCATTAG
QY	742	CTAGGCGTATTGCTAAAGTGGAACTGCTCTCGTCTCTTCTTCGGACCAACCATCTGGTG	801	742	CTAGGCGTATTGCTAAAGTGGAACTGCTCTCGTCTCTTCTTCGGACCAACCATCTGGTG
DB	901	CCCATAGAATCGCTGCTGTTGGAACCTGATTTCACGACCACTGTACACGATGAATCAGGAG	960	901	CCCATAGAATCGCTGCTGTTGGAACCTGATTTCACGACCACTGTACACGATGAATCAGGAG
QY	802	ATCCGTAGCAATGSCCTTTTCTAAGCTCAGCATGCTTGCACAGCAAAAGGGGTGATGCTT	861	802	ATCCGTAGCAATGSCCTTTTCTAAGCTCAGCATGCTTGCACAGCAAAAGGGGTGATGCTT
DB	961	ATGCATACCTCAGTGGCTTTTCAAAACGTCTCTCGGCTATTTGGGACGAGCGGTGAAGCAT	1020	961	ATGCATACCTCAGTGGCTTTTCAAAACGTCTCTCGGCTATTTGGGACGAGCGGTGAAGCAT
QY	862	ATGCAAAATGCAGATGTAAAGGGTTTCTCTGGAAGAGATTGCATGTAAACAGGTCATGATG	921	862	ATGCAAAATGCAGATGTAAAGGGTTTCTCTGGAAGAGATTGCATGTAAACAGGTCATGATG
DB	1021	ACACAAAACGCAATGCCAGAGTCTCCTTAGAAAATATTGCACAAAGCGTGGCTATAAAA	1080	1021	ACACAAAACGCAATGCCAGAGTCTCCTTAGAAAATATTGCACAAAGCGTGGCTATAAAA
QY	922	ATGTCCTTAAGCTGCACACTACTGATATTGCAATTTAGTCACTTATTAAGATCGAGAGCT	981	922	ATGTCCTTAAGCTGCACACTACTGATATTGCAATTTAGTCACTTATTAAGATCGAGAGCT
DB	1081	ATGTCCTCAGATCTCACACCAACTGAAATTTGTATCGAGGCCTTCGAGCAAGTTCTTGAGCT	1140	1081	ATGTCCTCAGATCTCACACCAACTGAAATTTGTATCGAGGCCTTCGAGCAAGTTCTTGAGCT
QY	982	T 982		982	T 982
DB	1141	T 1141		1141	T 1141
RESULT 7					

[illegible]

PR	23-AUG-1999;	99US-0149902P.	
PR	23-AUG-1999;	99US-0149930P.	
PR	25-AUG-1999;	99US-0150566P.	
PR	26-AUG-1999;	99US-0150884P.	
PR	27-AUG-1999;	99US-0151065P.	
PR	27-AUG-1999;	99US-0151066P.	
PR	27-AUG-1999;	99US-0151080P.	
PR	30-AUG-1999;	99US-0151303P.	
PR	31-AUG-1999;	99US-0151438P.	
PR	01-SEP-1999;	99US-0151930P.	
PR	07-SEP-1999;	99US-0152363P.	
PR	10-SEP-1999;	99US-0153070P.	
PR	13-SEP-1999;	99US-0153758P.	
PR	15-SEP-1999;	99US-0154018P.	
PR	16-SEP-1999;	99US-0154039P.	
PR	20-SEP-1999;	99US-0154779P.	
PR	22-SEP-1999;	99US-0155139P.	
PR	23-SEP-1999;	99US-0155486P.	
PR	24-SEP-1999;	99US-0155659P.	
PR	28-SEP-1999;	99US-0156458P.	
PR	29-SEP-1999;	99US-0156596P.	
PR	04-OCT-1999;	99US-0157117P.	
PR	05-OCT-1999;	99US-0157753P.	
PR	06-OCT-1999;	99US-0157865P.	
PR	07-OCT-1999;	99US-0158029P.	
PR	08-OCT-1999;	99US-0158232P.	
PR	12-OCT-1999;	99US-0158369P.	
PR	13-OCT-1999;	99US-0159293P.	
PR	13-OCT-1999;	99US-0159294P.	
PR	13-OCT-1999;	99US-0159295P.	
PR	14-OCT-1999;	99US-0159329P.	
PR	14-OCT-1999;	99US-0159330P.	
PR	14-OCT-1999;	99US-0159331P.	
PR	14-OCT-1999;	99US-0159637P.	
PR	18-OCT-1999;	99US-0159638P.	
PR	21-OCT-1999;	99US-0159584P.	
PR	21-OCT-1999;	99US-0160741P.	
PR	21-OCT-1999;	99US-0160767P.	
PR	21-OCT-1999;	99US-0160768P.	
PR	21-OCT-1999;	99US-0160770P.	
PR	21-OCT-1999;	99US-0160814P.	
PR	21-OCT-1999;	99US-0160815P.	
PR	22-OCT-1999;	99US-0160980P.	
PR	22-OCT-1999;	99US-0160981P.	
PR	22-OCT-1999;	99US-0160989P.	
PR	25-OCT-1999;	99US-0161404P.	
PR	25-OCT-1999;	99US-0161405P.	
PR	25-OCT-1999;	99US-0161406P.	
PR	26-OCT-1999;	99US-0161359P.	
PR	26-OCT-1999;	99US-0161360P.	
PR	26-OCT-1999;	99US-0161361P.	
PR	28-OCT-1999;	99US-0161920P.	
PR	28-OCT-1999;	99US-0161992P.	
PR	28-OCT-1999;	99US-0161993P.	
PR	29-OCT-1999;	99US-0162142P.	
Query Match			15.7%; Score 188.4; DB 3; Length 1095;
Best Local Similarity			57.6%; Pred. No. 1.3e-33;
Matches 377; Conservative			0; Mismatches 271; Indels 6; Gaps 2;
QY	334	CCGAGGTCATGAAACTGCACAACTCCGTTGACGAGCTCTCCTGTTGAAGAGAAAT	393
DB	255	CTGGAAGTCTCTTCAATCTCCATTTTGATGAAGAACACACAGATTTTGAAGCAGAAAAAG	314
QY	394	CAGAAGAAGTCTGTTCTTACTTGAACGGGAGGTGTTATTTACCTAGTAGGAATGATCGGTT	453
DB	315	CTGAAGAGGTAAACCGTATTTAAATGGACGATCGATATCTTGTGGTAGATGGTT	374
QY	454	CTGGAAAAAGTACTGTGGGGAAGATTATGTCGAAAGTCTTGGGTTATTCGTTCTTTGATA	513
DB	375	CCGGAAAAAGCACTAGGGAAGATTATGGCAAGATCGCTTGGTTATACATTTCTTTGATT	434
QY	514	GTGACAAGTTAGTGGAGCAAGCTGTTTGAATGCCATCAGTTGGCCAAATATTCAAGGTCC	573

DB	435	GIGACACTTTGATCGAGCAGGCTATGAAGGNACTCTCTGTAGCTGAGATATTTTGACATT	494
QY	574	ATAGTGAAGCCCTTTTCGGGATAATGAGAGTAGTGTCTTGAGAGATTGTCTCTCCATG-	632
DB	495	TCGGTGAGAGTGTCTTTCAGAGAAAAAGAGACTGAAGCGTTAAAGAAACTCTCTTTGATGT	554
QY	633	--CGACGATTAGTTGTTGCCACCGAGGTGGTGTCTTATCCGACCAATTAACCTGAGAT	690
DB	555	ACCACCAAGTTGTGTTTCAACCCGGGGAGGGCGAGTTATAAGACCCATCAATTGGAAGT	614
QY	691	ATATCAAGAGGGGCTATCTGTTTGTGTAGATGTCCCTTGATGCTCTTGTAGGCGTA	750
DB	615	ACATGCNTAAGGTATTAGTATTGGCTTGATACCTCTAGAAGCCCTTAGGCGCATAGAA	674
QY	751	TTGCTAAAGTGGGAACCTGCCTCTCGTCC---TCTTCTGGACCAACCATCTTGTGTATCCGT	807
DB	675	TAGCTGCTGTAGGAACGTGTTCAAGACCATGTCTACATGATGATGAGTCAGGGGACACAT	734
QY	808	ACGCAATGGCCTTTTCTAAGCTCAGCATGCTTGACAGCAAGGGGTGATGCTTATGCAA	867
DB	735	ACACAGCGGCTTTAAACCGTCTTTCAACGATTTGGGATGCGTGTGAAGCATACACTA	794
QY	868	ATGCAGATGTAAAGGTTTCTCTGGAAGAGATTGCATGTAAACAAGGTATGATGATGTCT	927
DB	795	AAGCCAGCGCAGAGATTTCCTTGGAGAATATTACTTTGAAGCTCGGTTATAGAAGTGTCT	854
QY	928	CTAAGCTGACACCTACTGATATTGCAATTGAGTGTCACTTCATTAAGATCGAGACT	981
DB	855	CAGATCTTACACCACTGAAATCGCCATTGAGGCTTTGAGCAAGTTCAGAGCT	908
RESULT 8			
AAC36746			
ID	AAC36746 standard; DNA; 1093 BP.		
AC	AAC36746;		
XX			
DT	17-OCT-2000 (first entry)		
XX			
DE	Arabidopsis thaliana DNA fragment SEQ ID NO: 14934.		
XX			
KW	Hybridisation assay; genetic mapping; gene expression control;		
KW	Protein identification; signal transduction pathway; metabolic pathway;		
KW	Promoter; termination sequence; ss.		
XX	Arabidopsis thaliana.		
OS	Arabidopsis thaliana.		
PN	EP1033405-A2.		
XX			
PD	06-SEP-2000.		
XX			
PF	25-FEB-2000; 2000EP-00301439.		
XX			
PR	25-FEB-1999; 99US-0121825P.		
PR	05-MAR-1999; 99US-012180P.		
PR	09-MAR-1999; 99US-0123548P.		
PR	23-MAR-1999; 99US-0125788P.		
PR	25-MAR-1999; 99US-0126264P.		
PR	29-MAR-1999; 99US-0126785P.		
PR	01-APR-1999; 99US-0127462P.		
PR	06-APR-1999; 99US-0128234P.		
PR	08-APR-1999; 99US-0128714P.		
PR	16-APR-1999; 99US-0129845P.		
PR	19-APR-1999; 99US-0130077P.		
PR	21-APR-1999; 99US-0130449P.		
PR	23-APR-1999; 99US-0130510P.		
PR	23-APR-1999; 99US-0130891P.		
PR	28-APR-1999; 99US-0131449P.		
PR	30-APR-1999; 99US-0132048P.		
PR	04-MAY-1999; 99US-0132407P.		
PR	05-MAY-1999; 99US-0132485P.		

[illegible]

Db 74 GRAAGTTAATGTCCAAAGTGCTGGTTATAGCTTCTTTGACTGCGACACTTTGATTGAAC 133
 QY 532 AGCTGTTGAATGCATCAGTTGCCCCAAATATTCAGGTCATAGTGAAGCTTCTTTC 591
 Db 134 AGCCGATGAATGGAATCTCTGTTGACAGATATTTGTTTCATCAGCGAGAGAAATTTTAA 193
 QY 592 GGGATAATGAGTAGTGTCTTGAGAGATTGTGCTCCATGCGACGATTAGTTGTGGCA 651
 Db 194 GAGGAAGAGACCGATGGCTTAAAGAGCTCTCTCGAGGTATCAAGTTGTGTTTCCA 253
 QY 652 CCGAGGTGTTGTTATCCGACCAATTAACATGGAGATATATGAAGAGGCGCTATCTG 711
 Db 254 CAGGTGAGGTGCAGTTATAAGACCCATTAACTGGAAGTATATGCATTAAGGAATCAGCA 313
 QY 712 TTTGTTAGATGGCC 727
 Db 314 TTTGGCTAGATGGCC 329

RESULT 13

ACN52580/c
 ID ACN52580 standard; cDNA; 573 BP.

AC ACN52580;

DT 02-DEC-2004 (first entry)

DE Cotton androecium tissue EST Clone ID: LIB3828-016-Q1-N6-D8, SEQ:7361.

XX Cotton; plant; EST; expressed sequence tag; transgenic plant; androecium;
 KW variety Nuotton33B; library LIB3828; molecular tag; molecular marker;
 KW genetic mapping; molecular mapping; seed germination; plant growth;
 KW plant quality; plant yield; plant breeding; tissue printing; ss.
 XX Gossypium hirsutum.

XX US2004123340-A1.

XX 24-JUN-2004.

XX 12-DEC-2001; 2001US-00021323.

XX 14-DEC-2000; 2000US-0255619P.

XX (DEIK/) DEIKMAN J.

XX (FENG/) FENG P C C.

XX (FINC/) FINCHER K L.

XX (ZIEG/) ZIEGLER T E.

XX Deikman J, Feng PCC, Fincher KL, Ziegler TE;

XX WPI; 2004-479808/45.

XX New isolated nucleic acid molecule that encodes a plant protein or its
 PT fragment, useful for isolating a variety of agronomically significant
 PT genes associated with plant growth, quality or yield, and as molecular
 PT tags to map genes.

XX Claim 1; SEQ ID NO 7361; 34pp; English.

XX The invention relates to 17880 cotton expressed sequence tags (ESTs;
 CC ACN45220-ACN63099). The ESTs were isolated from cDNA libraries generated
 CC from primed or non-primed seeds from variety DP50B, mature seeds from
 CC variety Coker 312 Boswell 96 Field, and androecium tissue, gynoeium
 CC tissue, developing fibres, carpal walls and septa from variety
 CC Nuotton33B. The invention also relates to substantially purified
 CC proteins or their fragments encoded by nucleic acid molecules of the
 CC invention, and to transformed plants having a nucleic acid construct
 CC comprising a nucleic acid of the invention. The cotton ESTs are useful as
 CC molecular tags to isolate genetic regions, to isolate genes, to map
 CC genes, to determine gene function and to determine whether genes are
 CC members of a particular gene family. The nucleic acid molecules may be
 CC used for isolating a variety of agronomically significant genes

CC associated with plant growth, quality, yield, and could also serve as
 CC links in metabolic and catabolic pathways. The nucleic acid molecules are
 CC also useful for identifying genes important in initiating and maintaining
 CC seed germination or that may be used to mitigate stresses encountered
 CC during seed germination. The ESTs additionally enable the acquisition of
 CC promoters and cis-regulatory elements which will be useful to express
 CC agronomically significant genes in these tissues and/or other tissues,
 CC and also permits the acquisition of molecular markers useful in breeding
 CC schemes, genetic and molecular mapping, and in cloning of agronomically
 CC significant genes. The nucleic acid molecules are further useful for
 CC detecting the expression level or pattern of a protein or mRNA and for
 CC detecting the presence or quantity of a protein by tissue printing. The
 CC present sequence represents a specifically claimed EST isolated from a
 CC cotton variety Nuotton33B androecium tissue cDNA library (LIB3828). The
 CC sequence data for this patent did not form part of the printed from the US
 CC specification, but was obtained in electronic format directly from the US
 CC patent office at seqdata.uspto.gov/sequence.html?docID=US20040123340
 XX
 SQ Sequence 573 BP; 183 A; 138 C; 86 G; 166 T; 0 U; 0 Other;

Query Match 11.3%; Score 135.2; DB 13; Length 573;
 Best Local Similarity 64.4%; Pred. No. 2.4e-21;
 Matches 219; Conservative 0; Mismatches 118; Indels 3; Gaps 1;

QY 628 CCATGCGACGATTAGTTGTCACCGAGGTGGTCTTATCCGACCAATTAACCTGA 687
 Db 573 CAAGAACAACACTGTTGTTCTACTGTGGAGTGCAGTTGTACTGGATGTAACCTGG 514
 QY 688 GATATAT---GAAGAGGGGCTTATCTGTTGGTTAGATGTCCTTGGATGCTCTTGCTA 744
 Db 513 ACTATATGCAGAAGAAGGGGATGTTGTTGGTTAGATGTACTCTTGAAGCCTTGGCAC 454
 QY 745 GCGTATTGCTAAAGTGGGAACCTGCTCTGCTCTTCTGACCAACCATCTGCTGATC 804
 Db 453 AAAGGATTGTCAGTAGTACTCATCTCTGCTCCCTTTGTCATTATGAAGATGGTGATC 394
 QY 805 CGTAGCGCAATGCGCTTTTCTAAGCTCAGCATGCTTGCAGCAAGGGGTGATGCTTATG 864
 Db 393 CATATACAAGGCTTTAAACGCTCTGTCTTACCTTTTGGAGCAGAGGGGTAAAAATTATG 334
 QY 865 CAAATGCAGATGTAAGGGTTTCTCTGGAAGAGATTGCATGTAAACAAGGTCAATGATG 924
 Db 333 CTAAAGCAAATGCCAGGGTTTCATTGTAAAGAAATGCTGGCAAACTAGGTTATAGAGATG 274
 QY 925 TCTCTAAGCTGACACCTACTCATATATGCAATTGAGTCACT 964
 Db 273 TATCAGATCTTACTCCACAGAGATCGCAATCGAGGCATT 234

RESULT 14

ADR60835

ID ADR60835 standard; cDNA; 1010 BP.

XX ADR60835;

DT 02-DEC-2004 (first entry)

XX Cotton cDNA sequence, SEQ ID 1616.

XX Cotton; ss; plant; cold tolerance; growth rate; cell cycle pathway;
 KW drought tolerance; plant disease resistance; galactomannan; lignin;
 KW plant growth regulator; heat tolerance; herbicide tolerance;
 KW homologous recombination; extreme osmotic condition tolerance;
 KW pathogen resistance; pest resistance; yield; photosynthesis; seed oil;
 XX stress resistance.

OS Gossypium hirsutum.

XX US2004181830-A1.

XX 16-SEP-2004.

XX 29-JAN-2004; 2004US-00767795.

useful for improving plant cold tolerance, manipulating growth rate in plant cells by modification of the cell cycle pathway, improving plant drought tolerance, providing increased resistance to plant disease, plant heat galactomannan (or lignin or plant growth regulators), improving plant growth tolerance, improving plant tolerance to herbicides, increasing the rate of homologous recombination in plants, improving plant tolerance to extreme osmotic conditions or to pathogens or pests, improving yield by modification of photosynthesis, modifying seed oil or protein yield and/or content, improving yield by modification of carbohydrate, nitrogen or phosphorus use and/or uptake, or improving yield by providing improved plant growth and development under at least one stress condition. The polynucleotide and polypeptide may also be used in recombinant DNA constructs, in physical arrays of molecules, as plant breeding markers, or in computer-based storage and analysis systems. The present sequence is a Cotton plant cDNA of the invention. NOTE: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from USPTO at seqdata.uspto.gov/sequence.html?docID=20040181830. However only 6585 polynucleotide sequences were available, the remaining 5213 polynucleotides and all 58798 protein sequences were not present.

Query Match	10.5%;	Score 125.8;	DB 13;	Length 614;
Best Local Similarity	61.7%;	Pred. No. 3.7e-19;		
Matches 235;	Conservative 0;	Mismatches 142;	Indels 4;	Gaps 2
Qy	588	TTTCGGGATAATGAGAGTAGTGCTCTTGAGAGATTGTCTCCATCGCAGCAATTAGTTGTT	647	
Db	1	TTCCAGAAAGAGGAGACTGAGGTATTGCAGAGGCTCTCTTCAAGAAACAGCTTTGTTGTT	60	
Qy	648	GCCACCGGAGTGGTGCTGTTTATCCGACCAATTAACCTGGAGATATAT---GAAGAGGGGC	704	
Db	61	TCTACTGGCGGAGGTGCAGTTGTTACGGGATGTGAACCTGGACTATATGCAGAAAGAGGG	120	
Qy	705	CTATCTGTTTGGTTAGATGTGCCCTTGGATGCTCTTGTAAGCGTATTGCTAAAGTGGGA	764	
Db	121	GTTGTGTGCTGTTAGATGTACTTTGGAAGCCTTGGCACACAGGATTGCTCGCATGAGT	180	
Qy	765	ACTGCCTCTCGTCCCTCTTCTGGACCAACCATCTGGTGATCCGTAGCAATGGCCTTTTCT	824	
Db	181	ACTCATCTCGTCCCTTTTGGCATTTGAACATGGCGATCCCTATACAAAGGCTTTAAAA	240	
Qy	825	AAGCTCACCATGCTTGCACGCAAGGGGTGATGCTTATGCAAAATGCAGATGAAGGTT	884	
Db	241	CGGCTGTCTTACCTTTTGGAGCTGAGGGGTAAAAATTATGCTAAAGCAAAATGCCCGGTT	300	
Qy	885	TCTCTG-CAAGAGATTGCATGTAAACAAGGTCATGATGATGTCTCTTAAGCTGCACACCTAC	943	
Db	301	TCATTGTAAGAAATTGCTGGCTTACTAGGTTATTGAGATGTATCAGATCTTACTCCAC	360	
Qy	944	TGATATTCGAATTGAGTCACT	964	
Db	361	AGAGATCGCAATCGAGGCATT	381	

Search completed: August 25, 2005, 17:38:21
Job time : 732 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 25, 2005, 16:54:33 ; Search time 367 Seconds
(without alignments)
5350.224 Million cell updates/sec

Title: US-10-660-226-9
Perfect score: 1200
Sequence: 1 ccgccaccagctacctgccc.....aaaaaaaaaaaaaaaaaaaaa 1200

Scoring table: IDENTITY NUC
Gapop 10_0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents_NA.*
1: /cgn2_6/prodata/1/ina/5A_COMB.seq.*
2: /cgn2_6/prodata/1/ina/5B_COMB.seq.*
3: /cgn2_6/prodata/1/ina/6A_COMB.seq.*
4: /cgn2_6/prodata/1/ina/6B_COMB.seq.*
5: /cgn2_6/prodata/1/ina/PCUS_COMB.seq.*
6: /cgn2_6/prodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	55	4.6	603	US-09-328-352-2023	Sequence 2023, Ap
2	53.8	4.5	1733	US-09-073-569-1	Sequence 1, Appli
3	53	4.4	1154	US-08-651-136C-7	Sequence 7, Appli
4	53	4.4	1154	US-09-229-911A-7	Sequence 7, Appli
5	52.8	4.4	222	US-08-481-190-15	Sequence 15, Appli
6	52.8	4.4	222	PCT-US93-00869-15	Sequence 15, Appli
7	52.6	4.4	1308	US-10-151-832-1	Sequence 1, Appli
8	52.6	4.4	1447	US-09-443-041A-27	Sequence 27, Appli
9	52.6	4.4	2045	US-09-152-060-22	Sequence 22, Appli
10	52.6	4.4	3871	US-08-599-455B-3	Sequence 3, Appli
11	52.6	4.4	3871	US-09-069-781B-3	Sequence 3, Appli
12	52.6	4.4	3871	US-09-137-132-3	Sequence 3, Appli
13	52.6	4.4	3871	US-08-864-564A-3	Sequence 3, Appli
14	52.6	4.4	3871	US-09-094-410-3	Sequence 3, Appli
15	52.6	4.4	3871	US-08-708-123D-3	Sequence 3, Appli
16	52.6	4.4	3871	US-08-583-153A-3	Sequence 3, Appli
17	52.6	4.4	3871	US-08-570-142D-3	Sequence 3, Appli
18	52.6	4.4	3871	US-08-638-524B-3	Sequence 3, Appli
19	52.4	4.4	1509	US-09-149-476-179	Sequence 179, App
20	52.2	4.4	2187	US-09-127-219B-2	Sequence 2, Appli
21	52.2	4.4	2267	US-09-917-265A-107	Sequence 107, App
22	52.2	4.4	2267	US-09-917-265A-109	Sequence 109, App
23	52	4.3	153	US-09-621-976-18058	Sequence 18058, A
24	52	4.3	160	US-09-621-976-18071	Sequence 18071, A
25	52	4.3	162	US-09-621-976-18068	Sequence 18068, A
26	52	4.3	179	US-09-621-976-18054	Sequence 18054, A
27	52	4.3	396	US-09-640-173-10	Sequence 10, Appli

c	28	52	4.3	396	4	US-09-713-550-10	Sequence 10, Appl
c	29	52	4.3	396	4	US-09-825-294-10	Sequence 10, Appl
c	30	52	4.3	396	4	US-09-970-966-10	Sequence 10, Appl
	31	51.8	4.3	1100	3	US-07-861-458C-4	Sequence 4, Appli
	32	51.6	4.3	1225	1	US-08-286-020-1	Sequence 1, Appli
	33	51.6	4.3	1225	1	US-08-603-919-1	Sequence 1, Appli
	34	51.6	4.3	1530	4	US-09-811-361-19	Sequence 19, Appli
	35	51.4	4.3	1512	2	US-08-909-965C-8	Sequence 8, Appli
	36	51.4	4.3	3602	3	US-09-402-929-1	Sequence 1, Appli
	37	51.2	4.3	160	4	US-09-621-976-18066	Sequence 18066, A
	38	51.2	4.3	1297	4	US-09-800-729-80	Sequence 80, Appl
c	39	51.2	4.3	3138	1	US-07-867-106-4	Sequence 4, Appli
	40	50.8	4.2	351	4	US-09-621-976-15134	Sequence 15134, A
	41	50.8	4.2	976	2	US-08-504-459-9	Sequence 9, Appli
	42	50.8	4.2	2080	4	US-10-003-392-1	Sequence 1, Appli
	43	50.8	4.2	6065	4	US-09-800-729-35	Sequence 35, Appli
	44	50.8	4.2	28257	4	US-09-949-016-13076	Sequence 13076, A
	45	50.6	4.2	1414	4	US-09-501-115-5	Sequence 5, Appli

ALIGNMENTS

RESULT 1
US-09-328-352-2023
; Sequence 2023, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 2023
; LENGTH: 603
; TYPE: DNA
; ORGANISM: Acinetobacter baumannii
US-09-328-352-2023

Query Match 4.6%; Score 55; DB 4; Length 603;
Best Local Similarity 52.8%; Pred. No. 0.00012;
Matches 143; Conservative 0; Mismatches 125; Indels 3; Gaps 1;

QY	428	TATTTACCTAGTACGAAATGATGGTTCTGGAAGAAAGTACTGTGGGAAGATTATGCTGA	487
Db	93	TATTTATTTGGTAGGCCCATGGGGCAGGAAACAAACCGTTGGACGTCATTTAGCAGA	152
QY	488	AGTCTTGGGTTATTTCGTTCTTTTGATAGTACAAAGTTAGTGAGCAAGCTGTTGGAATGCC	547
Db	153	ACTATTAGGGCGTGAATTTTATAGATAGTATGATGAAATTGAGCGCAAGACAGG---	209
QY	548	ATCAGTTGCCCAATATTCAGAGTCCATAGTGAAGCCCTTTTCGGGATATAGAGATAG	607
Db	210	CACATATTCCTGGATTTTGAAGAAAGAGAGAGTCGGCTTCCGTCGCGTGAACCTGT	269
QY	608	TGCTTTGAGAGATTGTCCTCATCGACGATTAGTCTTGCCACCGGAGGTGGTGCTGT	667
Db	270	CGTTTTAATAGCTTACTTCACGTAAGCAATAGTCTTGCCACAGGTGGCGTGCTAT	329
QY	668	TATCCGACCAATTAAGTGGAGATATATGAAG	698
Db	330	TACTCAGCGCCTAATCGGGAGTTTTTAAAG	360

RESULT 2
US-09-073-569-1
; Sequence 1, Application US/09073569
; Patent No. 6084088
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Grossmann, Angelika

```
;
; TITLE OF INVENTION: NOVEL TUMOR ANTIGENS
;
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ZymoGenetics, Inc.
; STREET: 1201 Eastlake Avenue East
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98102
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/073,569
; FILING DATE:
; CLASSIFICATION:
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Sawislak, Deborah A
; REGISTRATION NUMBER: 37,438
; REFERENCE/DOCKET NUMBER: 97-14
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-442-6672
; TELEFAX: 206-442-6678
; TELEX:
;
; INFORMATION FOR SEQ ID NO: 1:
;
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1733 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 34...1344
; OTHER INFORMATION:
;
; US-09-073-569-1
;
; Query Match 4.5%; Score 53.8; DB 3; Length 1733;
; Best Local Similarity 75.3%; Pred. No. 0.00045;
; Matches 67; Conservative 0; Mismatches 22; Indels 0; Gaps 0;
;
; QY 1112 AGGAATGGAAGAGGAGCTAATAATCCGAAGTGTGCCCTTGCTGAAAAA 1171
; | | | | | | | | | | | | | | | | | | | | | | | | | | | |
; Db 1561 AGGAGTGGCTGCATGTATCTGATAATACAGACCTGTCTCTCCCAAAAAA 1620
; | | | | | | | | | | | | | | | | | | | | | | | | | | | |
;
; QY 1172 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1200
; | | | | | | | | | | | | | | | | | | | | | | | | | | | |
; Db 1621 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1649
; | | | | | | | | | | | | | | | | | | | | | | | | | | | |
;
; RESULT 3
; US-08-651-136C-7
; Sequence 7, Application US/08651136C
; Patent No. 6001639
; GENERAL INFORMATION:
; APPLICANT: Schulein, Martin
; APPLICANT: Andersen, Lene N.
; APPLICANT: Lassen, Soren F.
; APPLICANT: Kauppinen, Markus S.
; APPLICANT: Lange, Lene
; APPLICANT: Nielsen, Ruby I.
; APPLICANT: Ihara, Michiko
; APPLICANT: Takagi, Shinobu
;
; TITLE OF INVENTION: No. 6001639el Endoglucanases
;
; NUMBER OF SEQUENCES: 109
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6001639o No. 6001639disk of No. 6001639th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
```

```
;
; STATE: New York
; COUNTRY: United States of America
; ZIP: 10174-6401
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/651,136C
; FILING DATE: 21-MAY-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Lambiris, Elias J.
; REGISTRATION NUMBER: 33,728
; REFERENCE/DOCKET NUMBER: 4366.200-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
;
; INFORMATION FOR SEQ ID NO: 7:
;
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1154 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 51..935
;
; US-08-651-136C-7
;
; Query Match 4.4%; Score 53; DB 3; Length 1154;
; Best Local Similarity 65.8%; Pred. No. 0.00059;
; Matches 77; Conservative 0; Mismatches 40; Indels 0; Gaps 0;
;
; QY 1084 ACATAGAGCATCGTTGAGTGTATTTCTAAAGGAATCGAAGAGGAGCTAATAATCCGAAG 1143
; | | | | | | | | | | | | | | | | | | | | | | | | | | | |
; Db 1025 ATATAGTCTTCGAGTACATACTATTGATGAATATAGAGCGGCTCGGACCATGAGCAG 1084
; | | | | | | | | | | | | | | | | | | | | | | | | | | | |
;
; QY 1144 TGTGCCGTGGCTGAAAAA 1200
; | | | | | | | | | | | | | | | | | | | | | | | | | | | |
; Db 1085 ATGCCATTGTATATAA 1141
; | | | | | | | | | | | | | | | | | | | | | | | | | | | |
;
; RESULT 4
; US-09-229-911A-7
; Sequence 7, Application US/09229911A
; Patent No. 6387690
; GENERAL INFORMATION:
; APPLICANT: Schulein, Martin
; APPLICANT: Andersen, Lene N.
; APPLICANT: Lassen, Soren F.
; APPLICANT: Kauppinen, Markus S.
; APPLICANT: Lange, Lene
; APPLICANT: Nielsen, Ruby I.
; APPLICANT: Ihara, Michiko
; APPLICANT: Takagi, Shinobu
;
; TITLE OF INVENTION: No. 6387690el Endoglucanases
;
; NUMBER OF SEQUENCES: 109
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6387690o No. 6387690disk of No. 6387690th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: United States of America
; ZIP: 10174-6401
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/229,911A
```

```

; FILING DATE: 13-Jan-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/651,136
; FILING DATE: 21-MAY-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Lambiris, Elias J.
; REGISTRATION NUMBER: 33,728
; REFERENCE/DOCKET NUMBER: 4366.200-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1154 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 51..935
; SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-09-229-911A-7

Query Match 4.4%; Score 53; DB 3; Length 1154;
Best Local Similarity 65.8%; Pred. No. 0.00059;
Matches 77; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

QY 1084 ACATAGAGCATCTTGAGTTATTGTTAAAGGAATGGAAGGAGCTAATAATCCGAG 1143
Db 1025 ATATAGTCTTCCGATACATACTATTGAATGAAATAAGAGCGCTCGGACCATGAGCAG 1084

QY 1144 TGTGCGCTTGGCTGAAAAA 1200
Db 1085 ATCCATTGATAAAAAA 1141
```

```

RESULT 5
US-08-481-190-15
; Sequence 15, Application US/08481190
; Patent No. 6160204
; GENERAL INFORMATION:
; APPLICANT: John C. Steffens
; TITLE OF INVENTION: Polyphenol Oxidase cDNA
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Yahwak & Associates
; STREET: 25 Skytop Drive
; CITY: Trumbull
; STATE: Connecticut
; COUNTRY: USA
; ZIP: 06611
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Microsoft Word 4.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/481,190
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION:
; APPLICATION NUMBER: 203,533
; FILING DATE: 02-24-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: George M. Yahwak
; REGISTRATION NUMBER: 26,824
; REFERENCE/DOCKET NUMBER: UA 816 CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (203)268-1951
; TELEFAX: (203)268-1951
; INFORMATION FOR SEQ ID NO: 15:
```

```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 222 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-481-190-15

Query Match 4.4%; Score 52.8; DB 3; Length 222;
Best Local Similarity 71.9%; Pred. No. 0.00028;
Matches 69; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

QY 1105 TTTGTAAGGAATGGAAGGAGCTAATAATCCGAAGTGCCTTGGCTGAAAAA 1164
Db 122 TTGTTAATGCTATGTATGAAATAAATAAATGTTTCCATTTTAAAAA 181

QY 1165 AAAAAA 1200
Db 182 AAAAAA 217

RESULT 6
PCT-US93-00869-15
; Sequence 15, Application PC/TUS9300869
; GENERAL INFORMATION:
; APPLICANT: John C. Steffens
; TITLE OF INVENTION: Polyphenol Oxidase cDNAs: Cloning
; TITLE OF INVENTION: and Applications
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Yahwak & Associates
; STREET: 25 Skytop Drive
; CITY: Trumbull
; STATE: Connecticut
; COUNTRY: USA
; ZIP: 06611
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Microsoft Word 4.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/00869
; FILING DATE: 19930129
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: George M. Yahwak
; REGISTRATION NUMBER: 26,824
; REFERENCE/DOCKET NUMBER: CRF D-1057
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (203)268-1951
; TELEFAX: (203)268-1951
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 222 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; PCT-US93-00869-15

Query Match 4.4%; Score 52.8; DB 5; Length 222;
Best Local Similarity 71.9%; Pred. No. 0.00028;
Matches 69; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

QY 1105 TTTGTAAGGAATGGAAGGAGCTAATAATCCGAAGTGCCTTGGCTGAAAAA 1164
Db 122 TTGTTAATGCTATGTATGAAATAAATAAATGTTTCCATTTTAAAAA 181

QY 1165 AAAAAA 1200
Db 182 AAAAAA 217
```



```
STATE: MA
COUNTRY: US
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows95
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/599,455B
FILING DATE: 22-JAN-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/583,153
FILING DATE: 28-DEC-1995
APPLICATION NUMBER: 08/570,142
FILING DATE: 11-DEC-1995
APPLICATION NUMBER: 08/569,485
FILING DATE: 08-DEC-1995
APPLICATION NUMBER: 08/566,622
FILING DATE: 04-DEC-1995
APPLICATION NUMBER: 08/562,663
FILING DATE: 27-NOV-1995
ATTORNEY/AGENT INFORMATION:
NAME: Meiklejohn, Ph.D., Anita L.
REGISTRATION NUMBER: 35,283
REFERENCE/DOCKET NUMBER: 07334/017001
TELEPHONE: 617-542-5070
TELEFAX: 617-542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 3871 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: Coding Sequence
LOCATION: 194...3688
US-08-599-455B-3

Query Match 4.4%; Score 52.6; DB 2; Length 3871;
Best Local Similarity 66.1%; Pred. No. 0.0014;
Matches 76; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

QY 1086 ATAGAGCATCGTTGAGTTATTTGTAAGGAATCGAAGAGGAGCTAATAATCCGAGTG 1145
DB 3744 ATAGATTATAGTTGTGGTGGGAGAGAGAAAGAACCCAGAGTCCAAATTTGAAATAAT 3803

QY 1146 TGCCGTTGGCTGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 1200
DB 3804 TGTTCCCACTGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 3858

RESULT 11
US-09-069-781B-3
Sequence 3, Application US/09069781B
Patent No. 6287782
GENERAL INFORMATION:
APPLICANT: Tartaglia, Louis A.
APPLICANT: Tepper, Robert I.
APPLICANT: Culpepper, Janice A.
APPLICANT: White, David W.
TITLE OF INVENTION: THE OB RECEPTOR AND METHODS FOR
TITLE OF INVENTION: THE DIAGNOSIS AND TREATMENT OF BODY WEIGHT DISORDERS,
TITLE OF INVENTION: INCLUDING OBESITY AND CACHEXIA
NUMBER OF SEQUENCES: 50
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson, P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA

STATE: MA
COUNTRY: US
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows95
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/599,781B
FILING DATE: 29-APRIL-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/864,564
FILING DATE: 28-MAY-1997
APPLICATION NUMBER: US 08/708,123
FILING DATE: 03-SEP-1996
APPLICATION NUMBER: US 08/638,524
FILING DATE: 26-APR-1996
APPLICATION NUMBER: US 08/599,455
FILING DATE: 22-JAN-1996
APPLICATION NUMBER: US 08/583,153
FILING DATE: 28-DEC-1995
APPLICATION NUMBER: US 08/570,142
FILING DATE: 11-DEC-1995
APPLICATION NUMBER: US 08/569,485
FILING DATE: 08-DEC-1995
APPLICATION NUMBER: US 08/566,622
FILING DATE: 04-DEC-1995
APPLICATION NUMBER: US 08/562,663
FILING DATE: 27-NOV-1995
ATTORNEY/AGENT INFORMATION:
NAME: Meiklejohn, Ph.D., Anita L.
REGISTRATION NUMBER: 35,283
REFERENCE/DOCKET NUMBER: 07334/082001
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 3871 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: Coding Sequence
LOCATION: 194...3688
US-09-069-781B-3

Query Match 4.4%; Score 52.6; DB 3; Length 3871;
Best Local Similarity 66.1%; Pred. No. 0.0014;
Matches 76; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

QY 1086 ATAGAGCATCGTTGAGTTATTTGTAAGGAATCGAAGAGGAGCTAATAATCCGAGTG 1145
DB 3744 ATAGATTATAGTTGTGGTGGGAGAGAGAAAGAACCCAGAGTCCAAATTTGAAATAAT 3803

QY 1146 TGCCGTTGGCTGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 1200
DB 3804 TGTTCCCACTGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 3858

RESULT 12
US-09-137-132-3
Sequence 3, Application US/09137132
Patent No. 6380363
GENERAL INFORMATION:
APPLICANT: Tartaglia, Louis A.
APPLICANT: Tepper, Robert I.
APPLICANT: Culpepper, Janice A.
APPLICANT: White, David W.
TITLE OF INVENTION: THE OB RECEPTOR AND METHODS FOR
TITLE OF INVENTION: THE DIAGNOSIS AND TREATMENT OF BODY WEIGHT DISORDERS,
TITLE OF INVENTION: THE DIAGNOSIS AND TREATMENT OF OBESITY AND CACHEXIA
NUMBER OF SEQUENCES: 50
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson, P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
```

;/ TITLE OF INVENTION: INCLUDING OBESITY AND CACHEXIA
;/ NUMBER OF SEQUENCES: 50
;/ CORRESPONDENCE ADDRESS:
;/ ADDRESSEE: Fish & Richardson, P.C.
;/ STREET: 225 Franklin Street
;/ CITY: Boston
;/ STATE: MA
;/ COUNTRY: US
;/ ZIP: 02110-2804
;/ COMPUTER READABLE FORM:
;/ MEDIUM TYPE: Diskette
;/ COMPUTER: IBM Compatible
;/ OPERATING SYSTEM: Windows95
;/ SOFTWARE: FastSEQ for Windows Version 2.0
;/ CURRENT APPLICATION DATA:
;/ APPLICATION NUMBER: US/09/137,132
;/ FILING DATE: 18-AUG-1998
;/ PRIOR APPLICATION DATA:
;/ APPLICATION NUMBER: 08/864,564
;/ FILING DATE: 28-MAY-1997
;/ APPLICATION NUMBER: 08/708,123
;/ FILING DATE: 03-SEP-1996
;/ APPLICATION NUMBER: 08/638,524
;/ FILING DATE: 26-APR-1996
;/ APPLICATION NUMBER: 08/599,455
;/ FILING DATE: 22-JAN-1996
;/ APPLICATION NUMBER: 08/583,153
;/ FILING DATE: 28-DEC-1995
;/ APPLICATION NUMBER: 08/570,142
;/ FILING DATE: 11-DEC-1995
;/ APPLICATION NUMBER: 08/569,485
;/ FILING DATE: 08-DEC-1995
;/ APPLICATION NUMBER: 08/566,622
;/ FILING DATE: 04-DEC-1995
;/ APPLICATION NUMBER: 08/562,663
;/ FILING DATE: 27-NOV-1995
;/ ATTORNEY/AGENT INFORMATION:
;/ NAME: Meiklejohn, Ph.D., Anita L.
;/ REGISTRATION NUMBER: 35,283
;/ REFERENCE/DOCKET NUMBER: 07334/019004
;/ TELECOMMUNICATION INFORMATION:
;/ TELEPHONE: 617-542-5070
;/ TELEFAX: 617-542-8906
;/ TELEX: 200154
;/ INFORMATION FOR SEQ ID NO: 3:
;/ SEQUENCE CHARACTERISTICS:
;/ LENGTH: 3871 base pairs
;/ TYPE: nucleic acid
;/ STRANDEDNESS: double
;/ TOPOLOGY: unknown
;/ MOLECULE TYPE: cDNA
;/ FEATURE:
;/ NAME/KEY: Coding Sequence
;/ LOCATION: 194...3688
;/ US-09-137-132-3

Query Match 4.4%; Score 52.6; DB 3; Length 3871;
Best Local Similarity 66.1%; Pred. No. 0.0014;
Matches 76; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

QY 1086 ATAGAGCATCGTTGAGTTATTTGTAAGGAAGGAGGAGCTAATAATCCGAAGTG 1145
Db 3744 ATAGATTATAGTTGTGGTGGGAGAGAGAAAGAACCCAGAGTCCAAATTTGAAATAAT 3803

QY 1146 TCCCGTTGGCTGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 1200
Db 3804 TGTTCCTCAACTGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 3858

RESULT 13
US-08-864-564A-3
; Sequence 3, Application US/08864564A
; Patent No. 6395498

;/ GENERAL INFORMATION:
;/ APPLICANT: Tartaglia, Louis A.
;/ APPLICANT: Tepper, Robert I.
;/ APPLICANT: Culpepper, Janice A.
;/ APPLICANT: White, David W.
;/ TITLE OF INVENTION: THE OB RECEPTOR AND METHODS FOR
;/ TITLE OF INVENTION: THE DIAGNOSIS AND TREATMENT OF BODY WEIGHT DISORDERS,
;/ TITLE OF INVENTION: INCLUDING OBESITY AND CACHEXIA
;/ NUMBER OF SEQUENCES: 50
;/ CORRESPONDENCE ADDRESS:
;/ ADDRESSEE: Fish & Richardson, P.C.
;/ STREET: 225 Franklin Street
;/ CITY: Boston
;/ STATE: MA
;/ COUNTRY: US
;/ ZIP: 02110-2804
;/ COMPUTER READABLE FORM:
;/ MEDIUM TYPE: Diskette
;/ COMPUTER: IBM Compatible
;/ OPERATING SYSTEM: Windows95
;/ SOFTWARE: FastSEQ for Windows Version 2.0
;/ CURRENT APPLICATION DATA:
;/ APPLICATION NUMBER: US/08/864,564A
;/ FILING DATE: 28-MAY-1997
;/ PRIOR APPLICATION DATA:
;/ APPLICATION NUMBER: 08/708,123
;/ FILING DATE: 03-SEP-1996
;/ APPLICATION NUMBER: 08/638,524
;/ FILING DATE: 26-APR-1996
;/ APPLICATION NUMBER: 08/599,455
;/ FILING DATE: 22-JAN-1996
;/ APPLICATION NUMBER: 08/583,153
;/ FILING DATE: 28-DEC-1995
;/ APPLICATION NUMBER: 08/570,142
;/ FILING DATE: 11-DEC-1995
;/ APPLICATION NUMBER: 08/569,485
;/ FILING DATE: 08-DEC-1995
;/ APPLICATION NUMBER: 08/566,622
;/ FILING DATE: 04-DEC-1995
;/ APPLICATION NUMBER: 08/562,663
;/ FILING DATE: 27-NOV-1995
;/ ATTORNEY/AGENT INFORMATION:
;/ NAME: Meiklejohn, Ph.D., Anita L.
;/ REGISTRATION NUMBER: 35,283
;/ REFERENCE/DOCKET NUMBER: 07334/019002
;/ TELECOMMUNICATION INFORMATION:
;/ TELEPHONE: 617-542-5070
;/ TELEFAX: 617-542-8906
;/ TELEX: 200154
;/ INFORMATION FOR SEQ ID NO: 3:
;/ SEQUENCE CHARACTERISTICS:
;/ LENGTH: 3871 base pairs
;/ TYPE: nucleic acid
;/ STRANDEDNESS: double
;/ TOPOLOGY: unknown
;/ MOLECULE TYPE: cDNA
;/ FEATURE:
;/ NAME/KEY: Coding Sequence
;/ LOCATION: 194...3688
;/ US-08-864-564A-3

Query Match 4.4%; Score 52.6; DB 3; Length 3871;
Best Local Similarity 66.1%; Pred. No. 0.0014;
Matches 76; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

QY 1086 ATAGAGCATCGTTGAGTTATTTGTAAGGAAGGAGGAGCTAATAATCCGAAGTG 1145
Db 3744 ATAGATTATAGTTGTGGTGGGAGAGAGAAAGAACCCAGAGTCCAAATTTGAAATAAT 3803

QY 1146 TCCCGTTGGCTGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 1200
Db 3804 TGTTCCTCAACTGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 3858

RESULT 14
US-09-094-410-3
; Sequence 3, Application US/09094410
; Patent No. 6403552
; GENERAL INFORMATION:
; APPLICANT: Tartaglia, Louis A.
; APPLICANT: Tepper, Robert I.
; APPLICANT: Culpepper, Janice A.
; APPLICANT: White, David W.
; TITLE OF INVENTION: THE OB RECEPTOR AND METHODS FOR
; TITLE OF INVENTION: THE DIAGNOSIS AND TREATMENT OF BODY WEIGHT DISORDERS,
; TITLE OF INVENTION: INCLUDING OBESITY AND CACHEXIA
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson, P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: US
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/094,410
; FILING DATE: 09-JUN-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/864,564
; FILING DATE: 28-MAY-1997
; APPLICATION NUMBER: 08/708,123
; FILING DATE: 03-SEP-1996
; APPLICATION NUMBER: 08/638,524
; FILING DATE: 26-APR-1996
; APPLICATION NUMBER: 08/599,455
; FILING DATE: 22-JAN-1996
; APPLICATION NUMBER: 08/583,153
; FILING DATE: 28-DEC-1995
; APPLICATION NUMBER: 08/570,142
; FILING DATE: 11-DEC-1995
; APPLICATION NUMBER: 08/566,622
; FILING DATE: 04-DEC-1995
; APPLICATION NUMBER: 08/562,663
; FILING DATE: 27-NOV-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Meiklejohn, Ph.D., Anita L.
; REGISTRATION NUMBER: 35,283
; REFERENCE/DOCKET NUMBER: 07334/019003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-542-5070
; TELEFAX: 617-542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3871 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: cdna
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 194...3688
US-09-094-410-3
Query Match 4.4%; Score 52.6; DB 3; Length 3871;
Best Local Similarity 66.1%; Pred. No. 0.0014;
Matches 76; Conservative 39; Indels 0; Gaps 0;

US-09-094-410-3
; Sequence 3, Application US/09094410
; Patent No. 6403552
; GENERAL INFORMATION:
; APPLICANT: Tartaglia, Louis A.
; APPLICANT: Tepper, Robert I.
; APPLICANT: Culpepper, Janice A.
; APPLICANT: White, David W.
; TITLE OF INVENTION: THE OB RECEPTOR AND METHODS FOR
; TITLE OF INVENTION: THE DIAGNOSIS AND TREATMENT OF BODY WEIGHT DISORDERS,
; TITLE OF INVENTION: INCLUDING OBESITY AND CACHEXIA
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson, P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: US
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/094,410
; FILING DATE: 09-JUN-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/864,564
; FILING DATE: 28-MAY-1997
; APPLICATION NUMBER: 08/708,123
; FILING DATE: 03-SEP-1996
; APPLICATION NUMBER: 08/638,524
; FILING DATE: 26-APR-1996
; APPLICATION NUMBER: 08/599,455
; FILING DATE: 22-JAN-1996
; APPLICATION NUMBER: 08/583,153
; FILING DATE: 28-DEC-1995
; APPLICATION NUMBER: 08/570,142
; FILING DATE: 11-DEC-1995
; APPLICATION NUMBER: 08/566,622
; FILING DATE: 04-DEC-1995
; APPLICATION NUMBER: 08/562,663
; FILING DATE: 27-NOV-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Meiklejohn, Ph.D., Anita L.
; REGISTRATION NUMBER: 35,283
; REFERENCE/DOCKET NUMBER: 07334/019003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-542-5070
; TELEFAX: 617-542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3871 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: cdna
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 194...3688
US-09-094-410-3
Query Match 4.4%; Score 52.6; DB 3; Length 3871;
Best Local Similarity 66.1%; Pred. No. 0.0014;
Matches 76; Conservative 39; Indels 0; Gaps 0;

US-08-708-123D-3
; Sequence 3, Application US/08708123D
; Patent No. 6482927
; GENERAL INFORMATION:
; APPLICANT: Tartaglia, Louis A.
; APPLICANT: Tepper, Robert I.
; APPLICANT: Culpepper, Janice A.
; APPLICANT: White, David W.
; TITLE OF INVENTION: THE OB RECEPTOR AND METHODS FOR
; TITLE OF INVENTION: THE DIAGNOSIS AND TREATMENT OF BODY WEIGHT DISORDERS,
; TITLE OF INVENTION: INCLUDING OBESITY AND CACHEXIA
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson, P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: US
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/708,123D
; FILING DATE: 03-SEP-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/638,524
; FILING DATE: 26-APR-1996
; APPLICATION NUMBER: 08/599,455
; FILING DATE: 22-JAN-1996
; APPLICATION NUMBER: 08/583,153
; FILING DATE: 28-DEC-1995
; APPLICATION NUMBER: 08/570,142
; FILING DATE: 11-DEC-1995
; APPLICATION NUMBER: 08/569,485
; FILING DATE: 08-DEC-1995
; APPLICATION NUMBER: 08/566,622
; FILING DATE: 04-DEC-1995
; APPLICATION NUMBER: 08/562,663
; FILING DATE: 27-NOV-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Meiklejohn, Ph.D., Anita L.
; REGISTRATION NUMBER: 35,283
; REFERENCE/DOCKET NUMBER: 07334/019001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-542-5070
; TELEFAX: 617-542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3871 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: cdna
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 194...3688
US-08-708-123D-3
Query Match 4.4%; Score 52.6; DB 4; Length 3871;
Best Local Similarity 66.1%; Pred. No. 0.0014;

Db 3744 ATAGATTATAGTTGTGGTGGAGAGAGAGAAACCCAGAGTCCANAATTTGAAAAATAAT 3803
Qy 1146 TGCCCGTTCGCTGAAAAA 1200
Db 3804 TGTCCCACTGAAAAA 3858

Matches	76;	Conservative	0;	Mismatches	39;	Indels	0;	Gaps	0;
Qy	1086	ATAGAGCATCGTTGAGTTATTGTAAGGAATGGAAGAGGAGGAGCTAATAATCCGAAGTG	1145						
Db	3744	ATAGATTATAGTTGTGGGTGGGAGAGAGAGAAAAACCAAGAGTCCAAATTTGAAAAATAAT	3803						
Qy	1146	TGCCGTTGGCTGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA	1200						
Db	3804	TGTTCCCAACTGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA	3858						

Search completed: August 25, 2005, 20:28:31
Job time : 368 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 25, 2005, 17:26:08 ; Search time 888 Seconds
(without alignments)

8842.013 Million cell updates/sec

Title: US-10-660-226-9

Perfect score: 1200

Sequence: 1 cgcgcaccagctacctgcc.....aaaaaaaaaaaaaaaaaaaa 1200

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 7331713 seqs, 3271544945 residues

Total number of hits satisfying chosen parameters: 14663426

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA.*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
- 10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
- 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
- 13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
- 14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
- 15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
- 16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
- 17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
- 18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq.*
- 19: /cgn2_6/ptodata/1/pubpna/US10G_PUBCOMB.seq.*
- 20: /cgn2_6/ptodata/1/pubpna/US10H_PUBCOMB.seq.*
- 21: /cgn2_6/ptodata/1/pubpna/US10I_PUBCOMB.seq.*
- 22: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
- 23: /cgn2_6/ptodata/1/pubpna/US11A_PUBCOMB.seq.*
- 24: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
- 25: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
- 26: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1200	100.0	1200	18	US-10-660-226-9 Sequence 9, Appli
2	1045.4	87.1	1059	18	US-10-425-114-30672 Sequence 30672, A
3	988.2	82.3	1474	20	US-10-425-115-16472 Sequence 16472, A
4	846.2	70.5	1228	18	US-10-425-114-35273 Sequence 35273, A
5	843.4	70.3	1583	20	US-10-425-115-16474 Sequence 16474, A
6	841.8	70.2	1241	18	US-10-425-114-31574 Sequence 31574, A
7	805	67.1	1273	19	US-10-767-701-13420 Sequence 13420, A

8	737.4	61.4	838	20	US-10-425-115-16473	Sequence 16473, A
9	660	55.0	1410	19	US-10-437-963-45320	Sequence 45320, A
10	574	47.8	1061	18	US-10-660-226-27	Sequence 27, Appl
11	370	30.8	1223	20	US-10-739-930-4532	Sequence 4532, Ap
12	367.8	30.6	1503	20	US-10-425-115-68871	Sequence 68871, A
13	364.4	30.4	1323	18	US-10-660-226-25	Sequence 25, Appl
14	363.4	30.3	899	18	US-10-660-226-11	Sequence 11, Appl
15	362.6	30.2	960	18	US-10-660-226-19	Sequence 19, Appl
16	352.6	29.4	738	20	US-10-425-115-16469	Sequence 16469, A
17	352.4	29.4	657	18	US-10-424-599-62654	Sequence 62654, A
18	342.2	28.5	1147	19	US-10-437-963-84942	Sequence 84942, A
19	335.2	27.9	536	18	US-10-425-114-6610	Sequence 6610, Ap
20	323.8	27.0	641	20	US-10-739-930-1680	Sequence 1680, Ap
21	293	24.4	397	19	US-10-437-963-84944	Sequence 84944, A
22	270	22.5	544	18	US-10-660-226-15	Sequence 15, Appl
23	241.6	20.1	1327	18	US-10-425-114-29872	Sequence 29872, A
24	241.6	20.1	1354	18	US-10-424-599-70058	Sequence 70058, A
25	233.6	19.5	637	21	US-10-487-901-57	Sequence 57, Appl
26	231.2	19.3	492	18	US-10-424-599-25738	Sequence 25738, A
27	227	18.9	1041	19	US-10-767-701-11266	Sequence 11266, A
28	222.6	18.6	1256	19	US-10-437-963-2340	Sequence 2340, Ap
29	206.2	17.2	365	20	US-10-425-115-151552	Sequence 151552, A
30	193.2	16.1	488	20	US-10-425-115-162636	Sequence 162636, A
31	193	16.1	1332	20	US-10-739-930-523	Sequence 523, App
32	191	15.9	967	18	US-10-424-599-137299	Sequence 137299, A
33	191	15.9	967	18	US-10-425-114-29612	Sequence 29612, A
34	190.4	15.9	524	18	US-10-425-114-18884	Sequence 18884, A
35	163	13.6	660	21	US-10-487-901-5640	Sequence 5640, Ap
36	162.2	13.5	407	20	US-10-425-115-158029	Sequence 158029, A
37	157	13.1	309	11	US-09-864-408A-8601	Sequence 8601, Ap
38	149.8	12.5	696	19	US-10-437-963-2317	Sequence 2317, Ap
39	149.2	12.4	676	20	US-10-425-115-145002	Sequence 145002, A
40	141.4	11.8	1096	18	US-10-424-599-11365	Sequence 11365, A
41	139.4	11.6	1183	18	US-10-660-226-21	Sequence 21, Appl
42	138.4	11.5	329	21	US-10-487-901-5638	Sequence 5638, Ap
43	137.4	11.5	513	18	US-10-424-599-19179	Sequence 19179, A
44	135.2	11.3	573	19	US-10-021-323-7361	Sequence 7361, Ap
45	131.2	10.9	332	20	US-10-425-115-33448	Sequence 33448, A

ALIGNMENTS

RESULT 1

US-10-660-226-9
; Sequence 9, Application US/10660226
; Publication No. US20040064848A1
; GENERAL INFORMATION:
; APPLICANT: Cahoon, Rebecca E.
; APPLICANT: Falco, S. Carl
; APPLICANT: Famodu, Layo O.
; APPLICANT: Hitz, William D.
; APPLICANT: Rendina, Alan
; TITLE OF INVENTION: Chorismate Biosynthesis Enzymes
; FILE REFERENCE: BB-1159-C
; CURRENT APPLICATION NUMBER: US/10/660,226
; CURRENT FILING DATE: 2003-09-11
; PRIOR APPLICATION NUMBER: US/09/354,501
; PRIOR FILING DATE: 1999-07-16
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/093,611
; PRIOR FILING DATE: EARLIER FILING DATE: July 21, 1998
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 9
; LENGTH: 1200
; TYPE: DNA
; ORGANISM: Zea mays
US-10-660-226-9

Query Match 100.0%; Score 1200; DB 18; Length 1200;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1200; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 656 AGTGTGCTGTTTATCCGACCAATTAACCTGGAGATATATGAAGAGGGGCTATCTGTTTG 715
Db 541 AGTGTGCTGTTTATCCGACCAATTAACCTGGAGATATATGAAGAGGGGCTATCTGTTTG 600
QY 716 GTTAGATGTGCCCTTCGATGCTCTTCTAGGCGTATTGCTAAAGTGGAACTGCCTCTCG 775
Db 601 GTTAGATGTGCCCTTCGATGCTCTTCTAGGCGTATTGCTAAAGTGGAACTGCCTCTCG 660
QY 776 TCCTCTTCTGGACCAACCAATCTGGTATCCGTACGCAATGGCCCTTTTCTAAGCTCAGCAT 835
Db 661 TCCTCTTCTGGACCAACCAATCTGGTATCCGTACGCAATGGCCCTTTTCTAAGCTCAGCAT 720
QY 836 GCTTGACACAAAGGGGTGATGCTTATGCAATGAGATGTAAGGGTTTCTCTGGAAGA 895
Db 721 GCTTGACACAAAGGGGTGATGCTTATGCAATGAGATGTAAGGGTTTCTCTGGAAGA 780
QY 896 GATTGATGTAAACAGGTCTATGATGCTCTTAAGCTGACACCTACTGATATTGCAAT 955
Db 781 GATTGATGTAAACAGGTCTATGATGCTCTTAAGCTGACACCTACTGATATTGCAAT 840
QY 956 TGAGTCACTTCAAGATCGAGAGCTTCGTATCGAGCACACTGCTGATAGTTCAAGTAG 1015
Db 841 TGAGTCACTTCAAGATCGAGAGCTTCGTATCGAGCACACTGCTGATAGTTCAAGTAG 900
QY 1016 CGACGCGCAAGCTGAGTCCAGATCCAGAGGATACAGACCTTGTAGAACCCTTAATCCCTT 1075
Db 901 CGACGCGCAAGCTGAGTCCAGATCCAGAGGATACAGACCTTGTAGAACCCTTAATCCCTT 960
QY 1076 TGTTCGCCACATAGACATCGTTGAGTTATTTGTAAGAAATGGAAGAGGAGCTTAATA 1135
Db 961 TGTTCGCCACATAGACATCGTTGAGTTATTTGTAAGAAATGGAAGAGGAGCTTAATA 1020
QY 1136 ATCCGAAGTGTGCGCTTGCTGAAAA 1162
Db 1021 ATCCGAAGTGTGCGCTTGCTGAAAA 1047

RESULT 3
US-10-425-115-16472
; Sequence 16472, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 16472
; LENGTH: 1474
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MPT4577_115020C.1
US-10-425-115-16472

Query Match 82.3%; Score 988.2; DB 20; Length 1474;
Best Local Similarity 93.8%; Pred. No. 4.9e-261;
Matches 1042; Conservative 0; Mismatches 63; Indels 6; Gaps 1;
QY 52 GGATCCTCAGAGATTCAGATTCAGTTGAGCTATAGCGGTAGCCGACTGGTCCGCG 111
Db 4 GGACCTTCGAGGCGCGCGGAGTCTGATCCGACTAGTCGCGCGCTCTCTTCCCGCG 63
QY 112 GTCCCTCTCGCTCCACCGCGGCGGAACATGGAGGGGGGGCGTCCGCTGGCG 171
Db 64 GCAGCAGCAGAGACACCGCGAGCGAGCAATGGAGCGGGGGCGTGGCGC 123
QY 172 TGCAGGCGCGGCGGGGCTTCGGCTCCAGCGCGCACCGGGCGGCGCTACAGGGCGCCCA 231

Db 124 TCACAGACGCGGCGCGGCTTCGGCTCCGACAGCGCGCGGCGCGCTTACAGTGCCTCA 183
QY 232 CCGGAGGCTGAGAGTCCGCTGACCGCGGGGACCTCGGTCGCTGTCGGGCTCGGGGT 291
Db 184 TCGGAGGCTGAGAGTCCGCTGACCGCGGGGACCTCGGTCGCTGTCGGGCTCGGGGT 243
QY 292 CCAAGCCGCTCCACCGCTCCGACTCCGTCGCGAAGAAATCGTCCGAGAGCTCATGAAAACT 351
Db 244 CCAAGCCGCTCCACCG - - - - - CTCCGTCGCAAGAAATCGTCCGAGAGTCAAGAACT 297
QY 352 CCGCAAACTCCGTTGACGAAGCTCCCTGTTGAAAGAAAAATCAGAAAGTTCTGTCTTCT 411
Db 298 TGCATAAACTCCGTTGACGAAGCACTCTCTGTTGAAAGAAAAATCAGAAAGTTCTGTCTTCT 357
QY 412 ACTTCGAACGGGAGGTGTTTACCTAGTAGGAATCATGCGTTCTCGAAAAAGTACTGTGCG 471
Db 358 ACTTAAACCGGAGGTGTTTACTTAGTGGGAATGATGGGTTCTGAAAAAGTACTGTGCG 417
QY 472 GGAAGATTATGTCTGAAGTCTTTGGGTTATTTCCGTTCTTTGATAGTGACAAAGTTAGTGGAGC 531
Db 418 GGAAGATCATGTCTGAAGTCTTTGGGTTATTTCCGTTCTTTGATAGTGACAAATTTAGTGGAGC 477
QY 532 AAGCTGTTGGAATGCAATCAGTTGCCAAATATTCAGGTCATATAGTGAAGCCTTCTTTCT 591
Db 478 AAGCTGTTGGAATGCAATCAGTTGCCAAATATTCAGGTCATATAGTGAAGCCTTCTTTCT 537
QY 592 GGGTAATGAGAGTAGTCTTGAGAGATTGTCCTCCATCGAGAGTTAGTTGTTGGCA 651
Db 538 GGGTAATGAGAGTAGTCTTGAGAGATTGTCCTCCATCGAGAGTTAGTTGTTGGCA 597
QY 652 CCGGAGGTGGTCTGTTATCCGACCAATTAACCTGGAGATATATGAAGAGGGGCTTATCTG 711
Db 598 CCGGAGGTGGTCTGTTATCCGACCAATTAACCTGGAGATATATGAAGAGGGGCTTATCTG 657
QY 712 TTTGTTAGATGTGCCCTTGGATGCTCTTGCTAGGCGTATTGCTAAAGTGGAACTGTCCT 771
Db 658 TTTGTTAGATGTGCCCTTGGATGCTCTTGCTAGGCGTATTGCTAAAGTGGAACTGTCCT 717
QY 772 CTCGCTCTTCTTGACCAACCATCTGGTGATCCGCTAGCGAATGGCCTTTCTTAAGCTCA 831
Db 718 CTCGCTCTTCTTGACCAACCATCTGGTGATCCGCTAGCGAATGGCCTTTCTTAAGCTCA 777
QY 832 GCATGCTTGACAGCAAAAGGGGTGATGCTTATGCAAAATGCAGATGTAAAGGGTTTCTCTGG 891
Db 778 GCATGCTTGACAGCAAAAGGGGTGATGCTTATGCAAAATGCAGATGTAAAGGGTTTCTCTGG 837
QY 892 AAGAGATTGCATGTAAACAAAGTCAATGATGTTCTCTAAAGCTGACACCTACTGATATTG 951
Db 838 AAGAGATTGCATGTAAACAAAGTCAATGATGTTCTCTAAAGCTGACACCTACTGATATTG 897
QY 952 CAAATTGAGTCACTTCATTAAGATCGAGAGCTTCGTCATCGAGCACACTGCTGATAGTTAG 1011
Db 898 CAAATTGAGTCACTTCATTAAGATCGAGAGCTTCGTCATCGAGCACACTGCTGATAGTTAG 957
QY 1012 CTAGCGACGCGCAAGCTGAGTCGAGATCCAGAGATACAGACCTTTGTAGAACCTTAAATC 1071
Db 958 CTAGCGACGCGCAAGCTGAGTCGAGATCCAGAGATACAGACCTTTGTAGAACCTTAAATC 1017
QY 1072 CTTTGTGTTGCCACATAGAGCATCGTTGAGTTATTGTAAGGAATGGAAGAGGAGGT 1131
Db 1018 CTTTGTGTTGCCACATAGAGCATCGTTGAGTTATTGTAAGGAATGGAAGAGGAGGT 1077
QY 1132 AATAATCCGAAGTGTGCCGTTGGCTGAAAA 1162
Db 1078 AATAATCCGAAGTGTGCCGTTGGCTGAAAA 1108

RESULT 4
US-10-425-114-35273
; Sequence 35273, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:

RESULT 5

Query Match	70.3%	Score 843.4;	DB 20;	Length 1583;
Best Local Similarity	88.0%;	Pred. No. 3.6e-221;		
Matches 944; Conservative	0;	Mismatches 121;	Indels 8;	Gaps 2;
Qy	15	CTGSCCTTCCTCTCTCTCTTTTACACCTCACCTCGGATCGCTCAGAGAGTCAG--AG	72	
Db	91	CTTGGGCCCTGGCCCTCCCTCTGTTTTTCTTAGACCTCCGGATCGCTCAGAGATTGAGTCGT	150	
Qy	73	ATTTCGAGTTGAGCTATAGGCGGTACCGCACTGGTGGCGGCTCCCTCTTCGGCTTCCACCCG	132	
Db	151	AGTCGTACCCGCACTGGTGGCGCGCTCCCTTTTCCCGCGCAGCAGCAGACCAACCCG	210	
Qy	133	CGCAGCGCAACAAATGGAGGCGGGGGGGCTCGGCTGGCGCTGAGGCGCGGGCGCGGGCT	192	
Db	211	GTGACCGAGCAATGGAGCGGGGGGGCTGGGCTTGGCGCTGCGACACGGGGCGCGGCGCT	270	
Qy	193	TCGGCTCCAGCCGCGCACCGGGGGCGGCTACAGCGCGCCACACCGGAGCGCTCAGAGTCGCTG	252	
Db	271	TCGGCTCCGGCGCAGCGCGGGGGGGCTACAGTCGCCCATCGGAGGCTCAGAGTCGCTG	330	
Qy	253	ACCGGGCGGACCTGGCGGTTCGCTGTGGCGGCTCGCGGGTCCAAAGCCCGTTCGACCGGTCC	312	

Db 331 AACCGCGGAGCTGCGGTTGCGGTTGCGGGTCCAGCCCGTACCG---- 386
Qy 313 GACTCCGTGCGAAGAAATCGTCGGAGGTATGAAATCTGCACAACTCGTTGACGAAG 372
Db 387 --CTCCGTGCGAAGAAATCATCGGAGGTATGAAATCTGCATTAATCTCGTTGACGAAG 444
Qy 373 CTCTCTGTTGAAGAGAAATCAGAGAAAGTCTGTTCTACTTTGAACGGGAGGTGATTT 432
Db 445 CTCTCTGTTGAAGAGAAATCAGAGAAAGTCTGTTCTACTTTAAACGGGAGGTGATTT 504
Qy 433 ACCTAGTAGGAATGAGGTTCTGGAAAAGTACTGTGGGAAGATATGTCGAAGTCT 492
Db 505 ACTTAGTGGGAATGAGGTTCTGGAAAAGTACTGTGGGAAGATCATGTCGAAGTCT 564
Qy 493 TGGGTTATTCGTTCTTTGATAGTGACAAAGTTAGTGGGAAGTCTGTGGAATGCCATCAG 552
Db 565 TGGGTTATTCGTTCTTTGATAGTGACAAATAGTGGGAAGTCTGTGGAATGCCATCAG 624
Qy 553 TTGCCAAATATTCAGAGTCCATAGTGAAGCCTTTCTTGGGAATATGAGAGTAGTGTCT 612
Db 625 TTGCTCAAAATATTCAGAGTCCATAGTGAAGCCTTTCTTGGGAATATGAGAGTAGTGTCT 684
Qy 613 TGAGAGATTTGCTCTCATCGAGAGTATGAGTGGTCCACCGAGGTGCTGTTATCC 672
Db 685 TGAGGAGTGTCTCTCATCGAGAGTATGAGTGGTCCACCGAGGTGCTGTTATCC 744
Qy 673 GACCAATTAATCTGAGAGTATGAGAGGGGCTTATCTGTTGTTAGATGTCCTCTGG 732
Db 745 GACCAGTTAACTGGAATATATGAGAGGGGCTTATCCGTTGTTAGATGTCCTCTGG 804
Qy 733 ATGCTCTTGTAGCGGTATGCTAAAGTGGGAACCTGCTCTCTCTTCTTGGAACCAAC 792
Db 805 ATGCTCTTGTAGCGGTATGCTAAAGTGGGAACCTGCTCTCTCTCTTCTTGGAACCAAC 864
Qy 793 CATCTGGTATCGTAGGCAATGGCCTTTCTTAAGCTCAGCATGCTTGACAGCAAGGG 852
Db 865 CGTCCGGTATCATACAAATGGCCTTTCTTAAGCTCAGCATGCTTGACAGCAAGGG 924
Qy 853 GTGATGCTTATGCAAAATGAGATGTAAAGGTTTCTTGGGAAGATGTCATGTAACAAG 912
Db 925 GTGATGCTTATGCAAAATGAGATGTAAAGGTTTCTTGGGAAGATGTCATGTAACAAG 984
Qy 913 GTCATGATGATGCTCTAAGCTGACACCTATGATATGCAATGAGTCACTTCAAGA 972
Db 985 GTCATGCGATGCTCTAAGCTGATCGGACTGATATCGCAATGAGTCACTTCAAGA 1044
Qy 973 TCGAGAGCTTCGTCATCGAGCACACTGCTGATAGTTTCAAGTACGAGCGGCAAGTGT 1032
Db 1045 TCGAGAGTTTCGTCATCGAGCACGCTGCTGATAATCCAGTACGAGCTCGCAAGTGT 1104
Qy 1033 CGCAGATCCAGAGATACAGACCTTGTAGAACCCTTAATCCCTTTGTTGCCAC 1085
Db 1105 CACAGATCCAAAGATACAGACCTTGTATATCTTAATCTCTGTTTGTATC 1157

RESULT 6

US-10-425-114-31574
; Sequence 31574, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 31574

; LENGTH: 1241
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMFLB73191D05_FLI
US-10-425-114-31574

Query Match

Best Local Similarity 70.2%; Score 841.8; DB 18; Length 1241;
Matches 943; Conservative 0; Mismatches 122; Indels 8; Gaps 2;

Qy 15 CTGCGCTTCTCTCTCTCTTTTACACCTCACTCCGATCGCTCAGAGTFCAG--AG 72
Db 30 CTGCGCCCTGCGCTCTCTCTTTTCTAGACCCCGCATCGCTCAGAGATTGAGTCGT 89
Qy 73 ATTCCAGTTGACTATAGCGGTAGCGACTGCTGCGCGGTCCTCTCGGCTCCACCCG 132
Db 90 AGTGTGATCCCGACTAGTCGCGCGCTCTCTTTTCCCGCGCAGCAGCAGCACCACCG 149
Qy 133 GCGAGCGAAATGAGGCGCGGGGCGTCTGCGCTCTGCGCTGCGAGCGCGGGCGCGGGCT 192
Db 150 GTGACCGAGCAATGAGGCGCGGGGCGTGGGCTCTGGGCTGCGAGCGCGGGCGGCGCT 209
Qy 193 TCGGCTCCAGCGGCAACCGGGCGGCTTACAGCGGCCACCGGAGCTTGAGAGTCCGTG 252
Db 210 TCGGCTCCGGCCAGCGCGGGCGGCTTACAGTCCGCCATCGGAGCTGAGAGTCCGTG 269
Qy 253 ACCCGCGGAGCTCGGGTCTGCTGCGGGTCTGCGGGTCCAAAGCCGTCGACCGCTCC 312
Db 270 AACCGCGGAGCTCGGGTCTGCGGGTCTGCGGGTCCAAAGCCGTCGTAACCG---- 325
Qy 313 GACTCCGTGCGAAGAAATCGTCCGAGGTGATGAAATCTGCACAACTCCGTTGACGAAG 372
Db 326 --CTCGTGGGAAGAAATCATCCGAGGTGATGAAATCTGATTAATCTCCGTTGACGAAG 383
Qy 373 CTCTCTGTTGAAGAAATCAGAAAGTCTGTTCTTCTTGAACGGGAGGTGATTT 432
Db 384 CTCTCTGTTGAAGAAATCAGAAAGTCTGTTCTTCTTGAACGGGAGGTGATTT 443
Qy 433 ACCTAGTAGGATGATGGGTTCTGGAAGAAAGTACTGTGGGAAGATATGTCGAAGTCT 492
Db 444 ACTTAGTGGGAATGATGGGTTCTGGAAGAAAGTACTGTGGGAAGATCATGTCGAAGTCT 503
Qy 493 TGGGTTATTCGTTCTTTTATAGTGACAAAGTTAGTGGGAAGTCTGTGGAAGTCCATCAG 552
Db 504 TGGGTTATTCGTTCTTTTATAGTGACAAATAGTGGGAAGTCTGTGGAATGCCCTTCA 563
Qy 553 TTGCCAAATATTCAGAGTCCATAGTGAAGCCTTCTTTGCGGATTAATGAGAGTAGTGTCT 612
Db 564 TTGCTCAAAATATTCAGAGTCCATAGTGAAGCCTTCTTTGCGGATTAATGAGAGTAGTGTCT 623
Qy 613 TGAGAGATTTGCTCTCCATGCGAGCAGTATGTTGTTGCCACCGGAGGTGCTGTTATCC 672
Db 624 TGAGGAGTCTGCTCTCCATGCGACATGATGTTGTTGCCACCGGAGGTGCTGTCATCC 683
Qy 673 GACCAATTAATGAGAGATATGAAAGGGGCTATCTGTTTGGTTAGATGTCCTCTGG 732
Db 684 GACCAATTAATGAGAGATATGAAAGGGGCTATCTGTTTGGTTAGATGTCCTCTGG 743
Qy 733 ATGCTCTTGTAGGCGTATGCTAAAGTGGGAATGCGCTCTCTCTCTTCTTGGACCAAC 792
Db 744 ATGCTCTTGTAGGCGTATGCTAAAGTGGGAATGCTCTCTCTCTCTTCTTGGACCAAC 803
Qy 793 CATCTGGTATCCGTTACGCAATGGCCTTTCTTAAGCTCAGCATGCTTGACAGCAAGGG 852
Db 804 CGTCCGGTATCCATACAAATGGCCTTTCTTAAGCTCAGCATGCTTGACAGCAAGGG 863
Qy 853 GTGATGCTTATGCAAAATGAGAGTGAAGGTTTCTCTGGAAGAGATTGATGTAACAAG 912
Db 864 GTGATGCTTATGCAAAATGAGAGTGAAGGTTTCTCTGGAAGAGATTGATGTAACAAG 923
Qy 913 GTCATGATGATGCTCTTAAGCTGACACCTTCTGATATGTCATGATTCGATTCATTAAGA 972

Db 166 GGTCCACCGCGGAGCAACATGAGCGGGGCGCTCGCGCTGCGCGCTCAGCGCG 225
Qy 182 GCGCGCGGGCTTCGGCTCAGCGCGGACCGCGGCGGCTTACAGGCGCCACCGGAGCCT 241
Db 226 GCGCGCGGGCTTCGGCTCAGCGCGGACCGCGGCGGCTTACAGGCGCCACCGGAGCCT 285
Qy 242 GAGAGTCGCTGACCCCGCGGACCTCGCGTTCGCTGCGGGCTCGCGGTCCAAAGCCGT 301
Db 286 GAGAGTCGCTGACCCCGCGGACCTCGCGTTCGCTGCGGGCTCGCGGTCCAAAGCCGT 345
Qy 302 CGCACGCTCCGACTCCGTCGAGAAATCGTCCGAGGTCTGAAATCTGCAACATC 361
Db 346 CGCACGCTCCGACTCCGTCGAGAAATCGTCCGAGGTCTGAAATCTGCAACATC 405
Qy 362 CGTTGACGAGCTCTCTCTGTTGAAGAGAAATCAGAAAGTTCCTGTTCTACTTGAACGG 421
Db 406 CGTTGACGAGCTCTCTCTGTTGAAGAGAAATCAGAAAGTTCCTGTTCTACTTGAACGG 465
Qy 422 GAGGTGTATTTACTAGTAGGAATGAGTGGTTCGGAAGATGACTGTCGGGAAGATTAT 481
Db 466 GAGGTGTATTTACTAGTAGGAATGAGTGGTTCGGAAGATGACTGTCGGGAAGATTAT 525
Qy 482 GTCTGAAGTCTGGGTATTCGTTCTTGTATGACAAAGTTAGTGGAGCAAGCTGTTGG 541
Db 526 GTCTGAAGTCTGGGTATTCGTTCTTGTATGACAAAGTTAGTGGAGCAAGCTGTTGG 585
Qy 542 AATGCCATCAGTTGCCCAATATTTCAAGGTCCATAGTGAAGCCTTCTTTCGGGATATGA 601
Db 586 AATGCCATCAGTTGCCCAATATTTCAAGGTCCATAGTGAAGCCTTCTTTCGGGATATGA 645
Qy 602 GAGTAGTCTTGAGAGATTTGCTCCATGCGACGATAGTTGTTGCCACCGAGGTGG 661
Db 646 GAGTAGTCTTGAGAGATTTGCTCCATGCGACGATAGTTGTTGCCACCGAGGTGG 705
Qy 662 TGCTGTATTCGACCAATTAACCTGGAGATATATGAAGGGGCTTACTGTTGGTTAGA 721
Db 706 TGCTGTATTCGACCAATTAACCTGGAGATATATGAAGGGGCTTACTGTTGGTTAGA 765
Qy 722 TGTGGCCTTGGATGCTCTGCTAGGCGTATGCTAAAGTGGGAACCTGCTCGTCTCT 781
Db 766 TGTGGCCTTGGATGCTCTGCTAGGCGTATGCTAAAGTGGGAACCTGCTCGTCTCT 824
Qy 782 TCTGGACCAAC 792
Db 825 TCTGGACCAAC 835

RESULT 9

US-10-437-963-45320
; Sequence 45320, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 45320
; LENGTH: 1410
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_48297C.1

US-10-437-963-45320
Query Match 55.0%; Score 660; DB 19; Length 1410;
Best Local Similarity 81.7%; Pred. No. 1e-170;
Matches 797; Conservative 0; Mismatches 155; Indels 24; Gaps 2;
Qy 149 GCGCGGGCGCTCGCGCTTCAGCGCGGCGGGCGGGCTTCGGCTCAGCGCGGA 208
Db 78 GAGAGCGGCGCTCGCGCTTCAGCTCGCGCGCGGGGTTTCGGCGGCTCGGACCG 137
Qy 209 CGGG-----GGCGGCTACAGCGGCCACCGCGAGCTCAGAGTGC 250
Db 138 CGCGCGAGCGGCTTCAGCGCGGCGGGCGGGATCGGAGCTTGAAGGTGCG 197
Qy 251 TGACCGCGCGGACCTCGCGCTGTCGCGGCTCGCGGCTCCAAAGCGCTCGCACCGCT 310
Db 198 TGAGCGGCGGCTGCGGAGGCGCTGTCGCGCTCGCGGCTCCAAAGCGCTCGCCCG-- 255
Qy 311 CGGACTCGTTCGGAAGAAATCGTCCGAGGTCTGAAACTCGCAAACTCCGTTGACGA 370
Db 256 ----CTCCGTGCAAGAAATCGTCCGAGGTCTGAAACATTTGCATAACTCGGTTGATGA 311
Qy 371 AGCTCTCTGTTGAAGAGAAATCAGAAAGTTCCTGTTCTACTTTGAAGCGGAGGTGAT 430
Db 312 AGCCCTCTTCTAAAGAGAAATCAGAAAGTTCCTTCTATTTGAATGAGCGGTGAT 371
Qy 431 TTACCTAGTAGGAATGATGGGTTCTGGAAGATGACTGTGGGAAAGATTTATGCTCAAAGT 490
Db 372 TTACCTAGTAGGAATGATGGGTTCTGGAAGATGACTGTGGGAAAGATCATGCTGAAAGT 431
Qy 491 CTTGGGTTATTCGTTCTTGTAGTACAAAGTTAGTGGAGCAAGCTGTTGGAATGCGATC 550
Db 432 TTTGGGTTATTCGTTCTTGTAGTACAAAGTTAGTGGAGCAAGCTGTTGGGATGCTTTC 491
Qy 551 AGTTCGCCCAATATTCAGGTCCTAGTAGGAGCTTCTTCGGGATATAGAGTAGTGT 610
Db 492 AGTCTCTCAAAATTTTCAAGGTTCTAGTAGGAGCTTCTTTAGGATATAGAGTAGTGT 551
Qy 611 CTTGAGAGATTTGCTCCTCCATCGACGATTTAGTTGTTGCCACCGGAGGTGGTGTAT 670
Db 552 CTTGAGGAGTTGCTCTCAATGAAGGATTTAGTTGTTGCTAGTGGGTGGTGTAT 611
Qy 671 CGGACCAATTAACCTGGAGATATGAAAGGGGCTTCTGTTGGTTAGATGTCGCTT 730
Db 612 CCGACCAAGTTAACTGGAAATACATGAAGAGGGCTTCTGTTGGTTGGATGTCGCTT 671
Qy 731 GGATGCTCTGCTAGCGGTATTGCTAAAGTGGGAAGTCTCTCTGCTCTTCTGGAACA 790
Db 672 GGACGCTCTGCTAGCGGTATTGCTAAAGTGGGAGTCTGCTCTCCGCTCTCTTCTAGATCA 731
Qy 791 ACCATCTGCTGATCCGTACGCAATGGCCTTTTCTAAGCTCAGCATGCTTGCAACAGAAAG 850
Db 732 ACCATCTGCTGATCCATACAAATGGCTTTTCTAAGCTCAGCATGCTCGCGGAGCAAG 791
Qy 851 GGGTATGCTTATGCAATGCAAGATGTAAGGGTTTCTCTGGAAGAGATGCAATGTAACA 910
Db 792 GGGCGATGCTTATGCAATGCTGATGTCGAGGGTTTCTCTTGAAGAGATGCAATCTAACA 851
Qy 911 AGGTGATGATGTCCTTAAGCTGACACCTTCTGATATGCAATTTAGTCACTTCTATAA 970
Db 852 GGGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 911
Qy 971 GATCGAGAGCTTTCGTCATCGAGCACTGCTGATAGTTTTCAGCTAGCGAGCGCAAGCTGA 1030
Db 912 GATCGAGAGCTTTCGTCATTCGATGATGATGATGATGATGATGATGATGATGATGAT 971
Qy 1031 GTCGAGAGATTCGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1090
Db 972 CTCAGCTGCTCAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1031
Qy 1091 GCATGTTGAGTTATT 1106
Db 1032 ACCTTGTAGCGCTGTT 1047

RESULT 10
US-10-660-226-27
; Sequence 27, Application US/10660226
; Publication No. US20040064848A1
; GENERAL INFORMATION:
; APPLICANT: Cahoon, Rebecca E.
; APPLICANT: Falco, S. Carl
; APPLICANT: Famodu, Layo O.
; APPLICANT: Hitz, William D.
; APPLICANT: Rendina, Alan
; TITLE OF INVENTION: Chorismate Biosynthesis Enzymes
; FILE REFERENCE: BB-1159-C
; CURRENT APPLICATION NUMBER: US/10/660,226
; CURRENT FILING DATE: 2003-09-11
; PRIOR APPLICATION NUMBER: US/09/354,501
; PRIOR FILING DATE: 1999-07-16
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/093,611
; PRIOR FILING DATE: EARLIER FILING DATE: July 21, 1998
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 27
; LENGTH: 1061
; TYPE: DNA
; ORGANISM: Triticum aestivum
US-10-660-226-27

Query Match 47.8%; Score 574; DB 18; Length 1061;
Best Local Similarity 81.5%; Pred. No. 4.1e-147;
Matches 693; Conservative 0; Mismatches 145; Indels 12; Gaps 2;

QY 227 GCCACCGGAGCTCAGAGTCGCTGACCCGGGGACCTCGCGTCGTCGTCGGGCTCG 286
Db 1 GCACGAGGTGAGCTTCGCTGTCAGTCATCTGTGGGTTCGCGCGCGCTCGCGCGG 60
QY 287 CGGGTCCAAAGCCGTCGACCGCTCGACTCCGTGCGAAGAAATCGTCGGAGGTCATGA 346
Db 61 CGGGCCAAAGCCGTCGTCGCTCCGCTCCG---CGCAAGAAATCGTCGGAGGAGTCATGA 117
QY 347 AAATCGCAACATCCGTTGACGAAGCTCTCTCTGTTGAAGAAATAATCAGAAAGATTCT 406
Db 118 GAATTTGCATACTCCGTTGACGATGCCCTCTGTTGAAGAAATAATCAGAAAGATTCT 177
QY 407 GTTCTACTTGAACGGAGGTGATTTTACCTAGTAGAATGATGGGTTCTGGAAAAAGTAC 466
Db 178 TTTCCAGTTGAACGGTCGCTGATCTACCTAGTTGGAATGATGGGTTCTGGGGAAAAAGTAC 237
QY 467 TGTGGGGAAGATTATCTCTGAAAGTCTTGGGTTATTCGTTCTTTGATAGTGACAAGTTAGT 526
Db 238 GGTGGGGAAGATCTTGCTGAAAGTTTGGGTTATTCATTTCTCGACAGTGATAAATTGGT 297
QY 527 GGAGCAAGCTGTTTGAATGCCATCAGTTGCCCAAAATATTTCAAGGTCATAGTGAAGCCCTT 586
Db 298 CGAACAAGCTGTTGGCATGCTTCAGTTGCTCAAATTTTCAAGGTTTCATAGTGAAGCCCTT 357
QY 587 CTTTCCGGATATGAGAGTAGTGTCTTGAGAGATTTGTCTCTCCATGCGACGATTAGTTGT 646
Db 358 CTTTCAGAGATAATGAGAGTAGTGTCTTGAGGATTTGTCTCTCAATGCGCGCATTAGTTGT 417
QY 647 TGCACCGGAGGTGTGTTATCCGACCAATTAACTCGAGATATATCAAGAGGGGCT 706
Db 418 TGCTACTGGAGGTGTGTTATCCGACAGTTAACTGGAAAAATATCAAGAGGGGCT 477
QY 707 ATCTGTTTGGTTAGATGTGCCCTTGGATGCTCTTGCTAGGCGTATTGCTAAAGTGGGAAC 766
Db 478 ATCTGTTTGGTTAGATGTGCCCTTGAAGCTCTTGCAAGGCGTATTGCTAAAGTGGGAC 537
QY 767 TGCCTCTGCTCTCTTGCAACCAACATCTGGTATCCGTAGCGAATGGCCCTTTTCTAA 826
Db 538 TGCCTCGCTCTCTTCTAGATCAACCATCCGGTGATCATACAAATGGCCCTTTTCGAA 597
QY 827 GCTCAGCATGCTTGACAGCAAGGGGTGATGCTTATGCAAAATGCAGATGTAAGGGTTTC 886

Db 598 ACTCAGCATGCTCGCGAGCAAGGGCGATGCTTATCAAAATGCTGATGTCAGAGTTTC 657
QY 887 TCTGGAAGAGATTGCATGTAAACAAGGTCATGATGATGCTCTTAAGCTGACACCTACTGA 946
Db 658 TCTCGAAGAGATCGCATCTAAGCTGGGTCAATGACGACGCTCTTAAGCTGACACCGATTGA 717
QY 947 TATTGCAATTGAGTCACCTTCAATAGATCGAGAGCTTCTCATCGAGCACACTGCTGATAG 1006
Db 718 TATTGCTCTCGAGTCGCTCCAAAGATCGAGAGCTTTGTGTCGGAAGACACCGCT- ---- 772
QY 1007 TTCAGCTAGCGACGCGCAAGCTGAGTCGCAGATCCAGAGGATACAGACTTTGTAGAACT 1066
Db 773 ----GTCCCGGACTCACAACGGAATCGCAATCTCAAAGGATGCATACCTTGTAGGATAT 828
QY 1067 TAATCCCTTT 1076
Db 829 GAATCCTTTT 838

RESULT 11
US-10-739-930-4532
; Sequence 4532, Application US/10739930
; Publication No. US20040216190A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH
; FILE REFERENCE: 38-21(53377)B
; CURRENT APPLICATION NUMBER: US/10/739,930
; CURRENT FILING DATE: 2003-12-18
; NUMBER OF SEQ ID NOS: 11088
; SEQ ID NO 4532
; LENGTH: 1223
; TYPE: DNA
; ORGANISM: Triticum aestivum
; FEATURE:
; OTHER INFORMATION: Clone ID: TRIAE-23APR03-CLUSTER25671_1
US-10-739-930-4532

Query Match 30.8%; Score 370; DB 20; Length 1223;
Best Local Similarity 64.3%; Pred. No. 6.1e-91;
Matches 606; Conservative 0; Mismatches 325; Indels 11; Gaps 3;

QY 144 ATGAGAGCGGGGGCGTCGGCTGCGCTGAGAGGCGGGCGGGCGGGCTTCGGCTCCAGC 203
Db 87 AGGCGCGCGGGCGCTGGCTCCGCCACAGTCGAGGCGCGGGCGGGCGGGCTTCGGCTGCA 146
QY 204 CGGCACCGGGCGGCTTACAGCGCGCCACCGGAGCCTGAGAGTCGCTGACCCGCGCGGA 263
Db 147 GCGGGCGCGCGGGGACCGGGCGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCTCGGG 206
QY 264 CCGCGGTCGCTGTGCGGGGCTCGGGGTCGCGGTCGCAAGCCGCTCGCACCGCTCCGCTCG 323
Db 207 GCGCGCGCGCGGGGCTTCGGGCTCCAGCGCGCTCGCGCGCGCGCGCGCGCGCGG 266
QY 324 AAGAAATCGTCGGAGGTCATGAAAACTCGCACAACTCGGTCGAGCAAGCTCTCTGTTG 383
Db 267 CCCTCGGAGGCGCGCGGAGGCGGCGGAGTCGCCACGCTCTGCTGCGCTCTGACCTCGAG 326
QY 384 AAGAAAAATCAGAAAGAGTTCTGTTCTTGAACGGGAGGTGATTTTACTAGTAGGA 443
Db 327 GTCCTCAGAGCTGC- ----CGGTCTCAATTGAATGGGCCCTCTAGCTACAGAGAGAG 380
QY 444 ATGA--TGGGTTCTGAAAAAGTACTGTGGGGAAGATTATGTCGAAAGCTTTGGGTTATT 501
Db 381 CTGAAGTGGGCCCTGTAAAAACACGGGCGTTGGATATTCGCCGTTGGAATTGGGGGTT 440
QY 502 CGTCTTTGATAGTACAAAGTAGTAGGAGCAAGCTGTGGATGCCATCAGTTGCCCAA 561
Db 441 CATACTTGACACTGATAGTTTGGTAGAACAGGCGAGTCGGAATGCTTCCGTTGCTCAA 500
QY 562 TATTCAAGGTCCTAGTAGTAAGCCTTCTTTTCGGGATAATGAGAGTAGTGTCTTGTAGAGATT 621

Db 501 TATTCAGGTTACAGTGAAGCATTTTTCAGAGACAGTGAAGTAGTGTCTTCAGAGATT 560
Qy 622 TGCTCCATCGACAGATTAGTTGTTGCCACCGAGGTGCTGCTGTTATCCGACCAATTA 681
Db 561 TATCCTCAATGCATCGCTAGTTGTTGCTACTAGAGCGGTGCTGTTATTCGACCAATT 620
Qy 682 ACTGGAGATATGAAGAGGGGCTATCTGTTTGGTTAGATGTCCTTGGATGCTCTTG 741
Db 621 ACTGGAGATATGAAGAAAGGTCTGTCTATCATGTTGGATGTGCCTTTGGATGCACTTG 680
Qy 742 CTAGGCGTATGCTAAGTGGGAACTGCCTCTGCTCTCTCTGGACCAACCATCTGGTG 801
Db 681 CGAAGCGCATGACAAAGTTGGAGCTGCTTCTCGGCCCTTCTAGATCAGCCATCTGCTG 740
Qy 802 ATCCGTACGCAATGGCCCTTTCTAAGCTCAGCATGCTTGCACAGCAAAAGGGGTGATGCTT 861
Db 741 ATCCATACACAGCGGCTTTCACGAACCTAGCGTCTCGCTGAGCAAAAGGGGATGCTT 800
Qy 862 ATGCAAAATGAGATGTAAGGTTTCTCTGGAAGAGATTGCAATTAAGCAAGGTCAATG 921
Db 801 ATGCAAAACCGCATGTAAGGTTTCTCTTGAAGAGCTTGCAGCTAAAAGGGTCAATG 860
Qy 922 ATGCTCTAGCTGACACCTACTGATATTCGAATTGAGTCACTCATAGATCGAGCT 981
Db 861 ATGCTCTCAGCTAAACCCCACTGATATTGCTGTCAGGGCCCTTACAAAGATTAAAGATT 920
Qy 982 TCGTCTACGAGCA--CACTGCTGATAGTTTCACTAGCGACGCGCAAGCTGAGTGCAGA 1038
Db 921 TTGCTACTGAGCAATTCATGCGCAGTGCGCCATTCGACGACTTATAAATTGGTGCCAG 980
Qy 1039 TCCAGAGGATACAGACCTTGTAGAACCTTAATCCCTTTTGT 1080
Db 981 GACGAACGATGCAATCTCTAGAAATTTTCAGCCCCCAATTGT 1022

RESULT 12

US-10-425-115-68871/c
; Sequence 68871, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 68871
; LENGTH: 1503
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_162804C.1
US-10-425-115-68871

Query Match 30.6%; Score 367.8; DB 20; Length 1503;
Best Local Similarity 67.3%; Pred. No. 2.7e-90;
Matches 518; Conservative 0; Mismatches 252; Indels 0; Gaps 0;
Qy 228 CCCACCGGAGCTGAGATCGCTGACCCGCGGGACCTGCGTCTGCGGGCTCGC 287
Db 1137 CCAGCGGAGGCTCGCGTGGCGGGACAGCCCGGAGGCTGTGCTGCGGGCC 1078
Qy 288 GGGTCCAAGCCCGTCGACCGCTCCGACTCCGTCGGAAGAAATCGTCCGAGGTCATGAA 347
Db 1077 GATACGCGGAGGCGCGGATCTGCGCTCGTGGCGCGCAATCTCAGGAACAGGA 1018
Qy 348 AACTCCGACAACTCCGTTGACGAAGCTCTCTGTTGAAGAGAAATTCAGAAAGTTCTG 407

Db 1017 AAGGTCACACTCTCTGCTGATGACGCTCTCATACTACAGCAAAAGCCCGAGGATGTTCTG 958
Qy 408 TTCTACTTGAACGGAGGTGATTATTACCTAGTAGGAATGATGGGTTCTCGAAAAAGTACT 467
Db 957 CTTTACTTGGATGGCGGTTGCGTTATCTTGTGGNAATGATGGGTTTCAGCAAAACTACA 898
Qy 468 GTGGGAAGATTATGTCGAAATCTTGGGTTTATTCGTTCTTTGATAGTCAAGTTAGTG 527
Db 897 GTTGGGAAGATACTCCGAAGTGTAGGTTATTCGTTCTTCGACAGTGATAAGTTGGTA 838
Qy 528 GAGCAAGCTGTTGGATGCGATCAGTGGCCAAATATCAAGGTCCATAGTGAAGCTTTC 587
Db 837 GAGAAGGCTGTTGGTATTTCATCTGTGTGAGATCTTTTCAGCTCCATAGCGAAACATT 778
Qy 588 TTTTGGGATAATGAGAGTAGTGTCTTGGAGATTTGTCTCATGGAGCATTTAGTTGTT 647
Db 777 TTCAGAGATTAAGAGTGAGGTCTCGAGGATCTGTCAATGATCGGTTGGTTGTT 718
Qy 648 GCCACGGAGGTGGTCTGTTATCCGACCAATTAATCGAGATATATGAAGAGGGGCTTA 707
Db 717 GCAACCGGAGGTGGTGCAGTGATCCGACCAATTCATGAGGTTTACATGAAGAAAGGCTG 658
Qy 708 TCTGTTTGGTTAGATGTCCTTGGATGCTCTTGTAGCGGTATTCGTAAGTGGGAACT 767
Db 657 ACCGTATGGTTAGATGTCCCACTGGATGCATTTGCAAGAAGAAATCGCTGCTGTAGGA 598
Qy 768 GCTCTCGTCTCTTCTGGACCAACCATCTGTGTGATCCGTACGCAATGCGCTTTTCTAAG 827
Db 597 GGTCTCGACCACTCTTGATCAGGAATCTGGTATCTCTTATGCAAAAGGCTTATGCAAAA 538
Qy 828 CTCAGCATGCTGCACAGCAAGGGGTGATGCTTATGCAAAATGCAAGATGTAAGGGTTCT 887
Db 537 CTTACATCATCTTTTGAGCAAGAAATGGACTCGTATGCTAATGCTGATGCCAGAGTTTCA 478
Qy 888 CTGGAAGAGATTGATGATGAACAAGTCAATGATGCTCTTAAGTGCACCTACTGAT 947
Db 477 CTTGAACATATTTGCAATTAACAAGGCCCAATATGATGTCACTATATCTTACCTAGTACC 418
Qy 948 ATTGCAATTGAGTCACTTTCATAGATCGAGAGCTTCTGTCATCGAGCACAC 997
Db 417 ATCGCATGAGGCATTGCTAAGATGGAAGTTTCTTACCGAGAGAC 368

RESULT 13

US-10-660-226-25
; Sequence 25, Application US/10660226
; Publication No. US20040064848A1
; GENERAL INFORMATION:
; APPLICANT: Cahoon, Rebecca E.
; APPLICANT: Falco, S. Carl
; APPLICANT: Famodu, Layo O.
; APPLICANT: Hitz, William D.
; APPLICANT: Rendina, Alan
; TITLE OF INVENTION: Chorismate Biosynthesis Enzymes
; FILE REFERENCE: BB-1159-C
; CURRENT APPLICATION NUMBER: US/10/660,226
; CURRENT FILING DATE: 2003-09-11
; PRIOR APPLICATION NUMBER: US/09/354,501
; PRIOR FILING DATE: 1999-07-16
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/093,611
; PRIOR FILING DATE: EARLIER FILING DATE: July 21, 1998
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 25
; LENGTH: 1323
; TYPE: DNA
; ORGANISM: Triticum aestivum
US-10-660-226-25

Query Match 30.4%; Score 364.4; DB 18; Length 1323;
Best Local Similarity 71.9%; Pred. No. 2.2e-89;
Matches 476; Conservative 0; Mismatches 186; Indels 0; Gaps 0;

QY 333 TCAGGAGGTCATGAAATCGCACAACTCGTTGACGAGCTCTCCTGTTGAAGAGAAA 392
Db 170 TCAGGAGGTCATGAAAGCCCACTATCTGCTGATGAGGCTCTCGTACTAAAGCAAAA 229
QY 393 TCAGAAAGAGTTCTGTTCTACTTTGAACGGGAGGTGTATTTACTAGTAGGAATGATGGGT 452
Db 230 GCAGAGGAGCTGCTCCCTTACCTGATGACCGCTGTGTTTATCTAGTTGGAATGATGGGT 289
QY 453 TCTGAAAAAGTACTGTGGGGAAGATATGCTGAAAGTCTTGGGTTTATTCGTTCTTTGAT 512
Db 290 TCGGCAAAACTACAGTTGCGGAAGATAATAGCTGGAAGTACTAGGCTATTCATTTCTTTGAC 349
QY 513 AGTGACAAGTTAGTGAGCAAGCTGTGGAATGCCATCAGTTGCCCAATATTCAGAGTC 572
Db 350 AGTGATAAGCTGTTGAGCAGTCTGTGGCAATACCGTCGCTGGCTGAGATTTTTCAGGTC 409
QY 573 CATAGTGAAGCCCTTCTTCGGGATAATGAGAGTGTCTTCGAGAGATTTGTCCCTCAATG 632
Db 410 CACAGTGAAGCATTTCTTCAGAGATAACGAGAGTGAGGTACTAAGGGATTTGTCTCAATG 469
QY 633 GCACGATTAGTTGTTGCCACCGAGGTGTGCTGTTATCCGACCAATTAACCTGGAGATAT 692
Db 470 CACCGATTAATTTGTTGCAACAGAGGTGTGCGGTGATACGACCAATCAATTTGGAGTTAT 529
QY 693 ATGAAGAGGGGCTATCTGTTGTTGTTAGATGTGCCCTTCGGATGCTCTTCGTAGGGGTATT 752
Db 530 ATGAAGAAAGACTCACTATTTGTTGTTAGATGTGCCATTTGAGCGCCCTTCGAAGAGGATT 589
QY 753 GCTMAAGTGGGAAGTCTCTGTCCTCTTCTGGACCAACCATCTGCTGATCCGTACGCA 812
Db 590 GCTGCGGTTGGTACTCGGTACGACGCCCTCTGTCATCAGGAATCTGCTGATCCTTATGCA 649
QY 813 ATGGCCCTTTCTMAAGCTAGCATGCTTGCAAGCAAGGGGTGATGCTTATGCAATGCA 872
Db 650 AAGGCTATGCCAACTTACAGCACTTTTGAACAAAGAAATGGAATTCATATGCTAATGCT 709
QY 873 GATGTAAGGTTTCTCTGGAAGAGATTGCAATGTAAACAAGGTTCATGATGCTCTTAAG 932
Db 710 GATGCCCGAGTTTCCCTTGAAATATTTGCATTCAAACAAGGACATGATGTAATGTA 769
QY 933 CTGACACCTACTGATATTGCAATTTAGTCACTTTCATAAGATCGAGAGCTTCGTCAFCGAG 992
Db 770 CTTACACCAAGTGCCATCGCTATTGAGGCATTGCTAAAGATGAGAGCTTCTTACTAGAG 829
QY 993 CA 994
Db 830 AA 831

RESULT 14

US-10-660-226-11
; Sequence 11, Application US/10660226
; Publication No. US2004006484A1
; GENERAL INFORMATION:

; APPLICANT: Cahoon, Rebecca E.
; APPLICANT: Falcio, S. Carl
; APPLICANT: Famodu, Layo O.
; APPLICANT: Hitz, William D.
; APPLICANT: Rendina, Alan
; TITLE OF INVENTION: Chorisimate Biosynthesis Enzymes
; FILE REFERENCE: BB-1159-C
; CURRENT APPLICATION NUMBER: US/10/660,226
; CURRENT FILING DATE: 2003-09-11
; PRIOR APPLICATION NUMBER: US/09/354,501
; PRIOR FILING DATE: 1999-07-16
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/093,611
; PRIOR FILING DATE: EARLIER FILING DATE: July 21, 1998
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 11
; LENGTH: 899
; TYPE: DNA
; ORGANISM: Zea mays

US-10-660-226-11

Query Match 30.3%; Score 363.4; DB 18; Length 899;
Best Local Similarity 71.0%; Pred. No. 3.4e-89;
Matches 481; Conservative 0; Mismatches 196; Indels 0; Gaps 0;

QY 321 GCGAAGAAATCGTCCGGAGGTCATGAAAACTCGCAAACTCGTGTGACGAAGCTCTCCTG 380
Db 5 GAGGCGCAATCTGCAAGTGGAAACAGGAAAGGTCCACTACTCTGCTGATGACGCTCTCAT 64
QY 381 TTGAAGAGAAATCAGAAAGAGTTCTGTCTACTTTGAACGGGAGGTGTATTTACCTAGTA 440
Db 65 CTCACGCAAAAGCCAGGATGTTCTGCTTACTTTGGATGGCCGTTCGGTGTATCTTGT 124
QY 441 GGAATGATGGGTTCTGGAAGAAAGTACTCTGCGGGAAGATATGCTGGAAGTCTTGGGTTAT 500
Db 125 GGAATGATGGGTTCTGGAAGAAAGTACTCTGCGGGAAGATATGCTGGAAGTGTAGGTTAT 184
QY 501 TCGTTCTTTGATAGTGACAAAGTTAGTGAGCAAGCTGTGGAATGCCATCAGTTGCCCAA 560
Db 185 TCGTTCTTCGACAGTATAGTTGTTGAGAAAGGCTGTGTTGATTTCACTCTGTTGCTGAG 244
QY 561 ATATTCAAGGTCATAGTGAAGCCCTTTCTCGGGAATATGAGAGTGTCTTTGAGAGAT 620
Db 245 ATCTTTGAGTCCATAGCGAAACATTTCTCAGAGATAATGAGAGTGAGTCTCGACGGAT 304
QY 621 TTGTTCTCCATGCGAGATTTAGTTGTTGCCACCGAGGTGGTGTGTTATCCGACCAATT 680
Db 305 CTGTCATCAATGCTATCGGTTGTTGCAACCGAGGTGGTGTGATCCGACCAATC 364
QY 681 AACTCGAGATATATCAAGAGGGGCTATCTGTTGTTGATGATGTCCTTGGATGCTCTT 740
Db 365 AATTGAGATTACATGAGAAAGGGCTGACGTATGATGTCCTGATGATGATGATGATGAT 424
QY 741 GCTAGCGGTATTGTAAAGTGGAACTGCTCTCTGCTCTCTGCTCTGGAACCAACCATCTGT 800
Db 425 GCAAGAAAGATCGTCTGTAGGAACCGCTCTGACCACTCTTGATCAGGAATCCGCT 484
QY 801 GATCCGTAGCAATGCGCTTTTCTTAAGCTCAGCATGCTTGACAGCAAGGGGTGATGCT 860
Db 485 GATCCTTATGCAAGGCTTATGCAAACTTACGTCACTTTTGGACAAAGAAATGACTCG 544
QY 861 TATGCAATGCAAGTGTAAAGGTTTCTCTGGAAGAGATTGCAATGTAACCAAGGTGATGAT 920
Db 545 TATGTAATGCTGATGCCAGATTTCACTTGAACATATTTGCAATTAACCAAGGCTAAT 604
QY 921 GATGCTCTAAGCTGACACCTACTCATATTTGCAATTTGATGCTCTTCAATAAGATCGAGGC 980
Db 605 GATGCTCACTATATTTACACCTAGTACCATTGCGCATTTGAGGCATTGCTAAAGATGGAAGT 664
QY 981 TTCGTCTATCGACACAC 997
Db 665 TTCTTACCGAGAAGAC 681

RESULT 15

US-10-660-226-19
; Sequence 19, Application US/10660226
; Publication No. US2004006484A1
; GENERAL INFORMATION:

; APPLICANT: Cahoon, Rebecca E.
; APPLICANT: Falcio, S. Carl
; APPLICANT: Famodu, Layo O.
; APPLICANT: Hitz, William D.
; APPLICANT: Rendina, Alan
; TITLE OF INVENTION: Chorisimate Biosynthesis Enzymes
; FILE REFERENCE: BB-1159-C
; CURRENT APPLICATION NUMBER: US/10/660,226
; CURRENT FILING DATE: 2003-09-11
; PRIOR APPLICATION NUMBER: US/09/354,501
; PRIOR FILING DATE: 1999-07-16
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/093,611
; PRIOR FILING DATE: EARLIER FILING DATE: July 21, 1998

; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 19
; LENGTH: 960
; TYPE: DNA
; ORGANISM: Sorghum
US-10-660-226-19

Query Match 30.2%; Score 362.6; DB 18; Length 960;
Best Local Similarity 71.6%; Pred. No. 5.8e-89;
Matches 476; Conservative 0; Mismatches 189; Indels 0; Gaps 0;

QY	333	TCCGGAGGTCATGAAGAACTCGACAACTCGGTTCGAGAGCTCTCTGTTGAAGAGAAA	392
Db	59		
QY	393	TCAGAGAGAGTTCTGTTCTACTTGAACGGGAGGTGATTACCTAGTAGGAATGATGGGT	452
Db	119		
QY	453	TCAGAGAGAGTTCTGTTCTACTTGAACGGGAGGTGATTACCTAGTAGGAATGATGGGT	512
Db	179		
QY	513	AGTGACAAGTTAGTGAGCAAGCTGTTGGAAAGCCATCAGTTGCCCAATATTCAGAGTC	572
Db	239		
QY	573	CATAGTGAAGCTTCTCGGGATATGAGAGTAGTGTCTTGAGAGATTGTCCTCATG	632
Db	299		
QY	633	CGACGATTAGTTGTCACCGGAGGTGGTGTCTGTTATCCGACCAATTAACCTGGAGATAT	692
Db	359		
QY	693	ATGAAGAGGGGCTATCTGTTGGTTAGATGTCCTTGGATGCTCTGCTAGGCGGTATT	752
Db	419		
QY	753	GCTAAAGTGGAACTGCCTCTGTCCTCTCTGAGCAACCACTCTGGTATCGTACGCA	812
Db	479		
QY	813	ATGGCTTTTCTAAGCTCAGCATGCTTGACAGCAAGGGGTGATGCTTATGCAATGCA	872
Db	539		
QY	873	GATGTAAGGGTTTCTCTGGAAGAGATTGCATGTAACAAAGGTGATGATGCTCTTAAG	932
Db	599		
QY	933	CTGACACCTACTGATATTCGAATTCAGTACCTTATAGATCGAGAGCTTCGTCATCGAG	992
Db	659		
QY	993	CACAC 997	
Db	719	AAGAC 723	

Search completed: August 25, 2005, 20:43:27
Job time : 890 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 25, 2005, 14:05:03 ; Search time 4422 Seconds
(without alignments)
10329.517 Million cell updates/sec

Title: US-10-660-226-9
Perfect score: 1200
Sequence: 1 ccgccaccagctaccctgcc.....aaaaaaaaaaaaaaaaaaaaa 1200

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- EST:*
- 1: gb_est1:*
 - 2: gb_est2:*
 - 3: gb_hic:*
 - 4: gb_est3:*
 - 5: gb_est4:*
 - 6: gb_est5:*
 - 7: gb_est6:*
 - 8: gb_gss1:*
 - 9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	824	68.7	1305	3	AY104682	Zea mays
2	619.8	51.6	804	7	CN147205	WOUN1_48
3	610.6	50.9	667	6	CA078453	SCRLAM100
4	567.6	47.3	672	6	CA146167	SCVPR1207
5	566	47.2	582	5	BQ577650	3524_1_39
6	566	47.2	825	6	CB684025	OSJNEf13C
7	549	45.8	549	7	COS24949	3530_1_16
8	547.8	45.6	619	6	CA452817	RP3A-3_H0
9	546.6	45.6	833	6	CB657302	OSJNEc11D
10	542	45.2	609	2	AW671996	LGI_353_A
11	540.8	45.1	620	6	CF033002	QCf11b08.
12	537.2	44.8	814	7	CF636791	zmrmw00_0
13	530.8	44.2	650	6	CA105686	SCJFHR1C0
14	524	43.7	728	7	CN144547	WOUN1_23
15	520.8	43.4	818	6	CB684026	OSJNEf13C
16	516.4	43.0	809	6	CB656619	OSJNEc11D
17	513	42.8	854	6	CB657303	OSJNEc121
18	510.4	42.5	775	6	CA182703	SCZST314
19	508.8	42.4	701	7	CN144633	WOUN1_23
20	506.2	42.2	733	6	CA452716	RP3A1_1
21	505.2	42.1	789	6	CB656559	OSJNEc11B
22	505	42.1	797	6	CB641679	OSJNEb01H
23	503	41.9	504	7	COS19743	3530_1_13
24	502.6	41.9	659	6	CA112411	SCEQDB106

C	25	502.4	41.9	789	6	CB656620	OSJNEc11D
	26	500.8	41.7	582	4	BG355018	BJ472228
	27	497.8	41.5	640	4	BJ472228	BJ472228
	28	496.8	41.4	689	4	BJ248029	BJ248029
	29	496.4	41.4	612	6	CA118037	SCBGR104
	30	495	41.2	642	6	CD670151	3529_1_12
	31	493.6	41.1	786	7	CK827364	zmrmw01_0
	32	492.4	41.0	586	7	COS24948	3530_1_16
	33	491.2	40.9	772	6	CB656558	OSJNEc11B
	34	490.2	40.8	601	1	AI795555	614009G01
	35	488.6	40.7	504	7	COS19742	3530_1_13
	36	484.2	40.3	669	6	CA248324	SCCCFL509
	37	483.8	40.3	604	2	BE593776	WSI_102_A
	38	482.4	38.1	914	4	BG301287	HVSMBO02
	39	455.8	38.0	670	7	CF633545	zmrmw00_0
	40	455.6	38.0	513	2	AW671997	LGI_353_A
	41	450.6	37.5	649	1	AV909303	AV909303
	42	450.4	37.4	649	7	CF489674	POLL_59_A
	43	448.2	37.4	619	4	BG241548	RHI22_51
	44	446.4	37.2	711	6	CA137983	SCEPRT204
	45	435.4	36.3	1066	7	CR286591	CR286591

ALIGNMENTS

RESULT 1
AY104682
LOCUS AY104682 1305 bp mRNA linear HTC 16-OCT-2002
DEFINITION Zea mays PCO108968 mRNA sequence.
ACCESSION AY104682
VERSION AY104682.1 GI:21207760
KEYWORDS HTC.
SOURCE Zea mays
ORGANISM Zea mays
REFERENCE 1 (bases 1 to 1305)
AUTHORS Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whiteitt,M.S.,
Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.
TITLE Maize Mapping Project/Dupont Consensus Sequences for Design of
Overgo Probes
JOURNAL Unpublished (2002)
REFERENCE 2 (bases 1 to 1305)
AUTHORS Coe,E.H.
TITLE Direct Submission
JOURNAL Submitted (25-APR-2002) Maize Mapping Project, University of
Missouri, Columbia, MO 65211, USA
COMMENT If you are interested in getting corresponding physical clones,
these are publicly available from ZmDB and may be found by BLAST
searching at MSL, maizemap.org; ZmDB, www.zmdb.iastate.edu; TIGR,
www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the
maize cDNA sequences is either Virginia Walbot, Stanford or Pat
Schnable, Iowa State, then clones may be requested from ZmDB:
www.zmdb.iastate.edu.

FEATURES

Location/Qualifiers
1..1305
/organism="Zea mays"
/mol_type="mRNA"
/db_xref="MaizeDB:636337"
/db_xref="taxon:4577"
/clone_lib="Maize Mapping Project/DuPont Cornsensus
Library"
/note="this sequence is part of a project of EST
assemblies resulting from the application of public
contigs to seed Dupont contigs; this resource was
assembled by Dupont as part of a collaboration for the
overgo addressing of BACs in conjunction with the Maize
Mapping Project"

ORIGIN

Query Match 68.7%; Score 824; DB 3; Length 1305;

Best Local Similarity 87.4%; Pred. No. 1.8e-188;
Matches 938; Conservative 0; Mismatches 126; Indels 9; Gaps 3;

Qy		15	CCTGCCTTCTCTCTCCTCTCTTTACACCTCACCTCCGGATCGCTCAGAGACTCAG--AG	72
Db		66	CCTGCGCCCTGCGCTCCCTCTGTCTTTCTAGACCCC CGGATCGCTCAGAGAATTGAGTCGT	125
Qy		73	ATTTCGAGTTTGAGCTATAGGCGTAGCGACTGGTCGCCGCGTCCCTCTCGGCTCCACCCG	132
Db		126	AGTCGTACCAGACTAGTCGCGGCTCTCTTTCCCGCGCAGCAGCAGCAGCAACCCG	185
Qy		133	GCGAGCGAAACAATGAGGCGGGGGCGTCCGCTTGGCGCTGCAAGCGCGCGCGCGCGCT	192
Db		186	GTACCGAGCAATGAGGCGGGGGCGTGGCGCTGCAGACGCGCGCGCGCGGCGCT	245
Qy		193	TGCGCTCCAGCCGGCACCGGGGGCGCTACAGGGGCCACCGGGAGCCTCAGAGTCGCTG	252
Db		246	TCGGCTCCGCCACGCGCGGGGGCGCTCAGTCGCCCATCGGAGGCTCAGAGTCGCTG	305
Qy		253	ACCGGGCGGACCTGCGGCTCGCTGTGCGGCTCGCGGCTCAAAGCCGTCGCAACCGCTCC	312
Db		306	AACCGGGCGAGCTGCGGTTGCGGTGCGGGTCCAAGCCGTCGTACCG-----	361
Qy		313	GACTCCGTGCGAAGAAATCGTCCGGAGGTCATGAATACTCCACAACTCCGTTGACGAAG	372
Db		362	--CTCCGTGCGAAGAAATCATCCGGAGGTCATGAATACTTCATAACTCCGTTGACGAAG	419
Qy		373	CTCTCTGTTGAAGAGAAAATCAGAGAAGTCTGTCTACTTCTTAAGCGGAGGCTGATTT	432
Db		420	CTCTCTGTTGAAGAGAAAAATCAGAGAAGTCTGTCTACTTAAACCGGAGGCTGATTT	479
Qy		433	ACCTAGTAGGAATGATCGGTTCTGGAAAAAGTACTGTGGGAAAGATPATCTCTGAAGTCT	492
Db		480	ACTTAGTGGAAATGATGGGTTCTGGAAAAAGTACTGTGGGAAAGATCATGTCTGAAGTCT	539
Qy		493	TGGGTTATTCTGTTCTTTGATAGTACAAAGTTAGTGGAGCAAGCTGTTGGAAATGCCATCAG	552
Db		540	TGGGTTATTCTGTTCTTTGATAGTGACAAATTAGTGGAGCAAGCTGTTGGAAATGCCITCAG	599
Qy		553	TTGCCAAATATTCAAGGTCATAGTGAAGCCTTCTTTCCGGATAATGAGAGTAGTGCT	612
Db		600	TTGCTCAAAATATTCAAAAGTTTACAGTGAAGCCTTCTTTCCGGATAATGAGAGTAGCGTCT	659
Qy		613	TGAGAGATTGTCCTCCATCGCAGATTAGTGTGGCCACCGAGGTGGTCTGTTATCC	672
Db		660	TGAGGATCTGTCTCCATCGGACGATTAGTTGTTGCCACCGA-GHTGNCTGTCTATCC	718
Qy		673	GACCAATTAACTGSAGATATATGAAGAGGGCCCTATCTGTTGGTTAGATGTGCCCTGG	732
Db		719	GACCAGTTAACTGSAATATATGAAGAAGGCGCTATCCGTTTGGTTAGATGTGCCCTTGG	778
Qy		733	ATGCTCTTGCTAGGCGTATTGCTAAAGTGGAACTGCTCTCGTCTCTTTCTGGACCAAC	792
Db		779	ATGCTCTTGCTAGGCGATTGCTAAAGTGGAAACCGCTTCTCGTCTCTTTCTGGACCAAC	838
Qy		793	CATCTGCTGATCCGTACGCAATGSCCTTTTCTAAGCTCAGCATGCTTTCACAGCAAGGG	852
Db		839	CGTCCGCTGATCCATACAAATGGCTTTTCTAAGCTCAGCATGCTTTCAGAGCAAGGG	898
Qy		853	GTGATGCTTATGCAAAATGCAGATGAAGGGTTTCTCTGGAAAGATTGTCATGTAAAACAG	912
Db		899	GTGATGCTTATGCAAAATGCAGATGAAGGGTTTCTCTGGAAAGATTGTCATGTAAAACAG	958
Qy		913	GTCAATGATGATGCTCTTAAGCTGACACTTACTGATATTGCAATTGAGTCACTTCATAGA	972
Db		959	GTCAATGATGATGCTCTTAAGCTGATGCGCATGTATTCGCAATTTGAGTCACTTCATAGA	1018
Qy		973	TCGAGAGCTTCGTCATCGACACACTGCTGATAGTTCAGCTAGCGACGCGCAAGCTCAGT	1032
Db		1019	TCGAGAGTTTCGTCATCGACACGCTGCTGATTAATCCAGCTTAGGCACTCGAAGCTGAGT	1078
Qy		1033	CGCAGATCCAGAGGATACAGACTTGTAGAACCTTAATCCCTTTGTTTGGCAC	1085

Db 1079 CACAGATCCAAAGGATACAGACCTTGTAATATCTTAATCCTTCTGTTTGTAC 1131

RESULT 2	CN147205	804 bp	mRNA	linear	EST 01-APR-2004
LOCUS	WOUND1_48_G03_b1_A002	Wounded leaves	Sorghum bicolor	cdna clone	
DEFINITION	WOUND1_48_G03_A002 3', mRNA sequence.				
ACCESSION	CN147205	1	GI:45987613		
VERSION	EST.				
KEYWORDS	Sorghum bicolor (sorghum)				
SOURCE	Sorghum bicolor				
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Sorghum.				
REFERENCE	1 (bases 1 to 804)				
AUTHORS	Cordonnier-Pratt, M.-M., Suzuki, Y., Sugano, S., Klein, R.R., Liang, C., Sun, F., Sullivan, R., Brady, J., Eastman, A., Miller, V., Gonzalez, M., Anfuoso, C., Chhabra, D., Johnson, H., Kanran, D. and Pratt, L.H.				
TITLE	A Sorghum EST database: mechanically damaged and methyl jasmonate-treated leaves				
JOURNAL	Unpublished (2003)				
COMMENT	Other_ESTs: WOUND1_48_G03_G1_A002 Contact: Cordonnier-Pratt MM Laboratory for Genomics and Bioinformatics The University of Georgia, Department of Plant Biology Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA Tel: 706 542 1860 Fax: 706 583 0210 Email: mmpratt@uga.edu Library constructed by Dr. Yutaka Suzuki and Dr. Sumio Sugano in the Human Genome Center, University of Tokyo Institute of Medical Science; plant material and RNA prepared at Texas A & M University; sequencing done in the Laboratory for Genomics and Bioinformatics, University of Georgia. Sequence ends have been trimmed to exclude vector and regions below Phred quality 16. Three-prime sequences are presented as their reverse complement and have been trimmed to exclude polyA. Seq primer: Sug3-14 (TAGTCTAGCGCCGCGACC) POLYA=Yes.				

FEATURES	Location/Qualifiers	source
	1..804	
	/organism="Sorghum bicolor"	
	/mol_type="mRNA"	
	/cultivar="BTx623"	
	/db_xref="taxon:4558"	
	/clone="WOUND1_48_G03_A002"	
	/lab_host="DH10B-T1 phage-resistant E. coli"	
	/clone_lbb="Wounded leaves"	
	/notes="Organ: Leaf; Vector: pME18S-FL3; Site_1: XhoI; Site_2: XhoI; The library was prepared from polyA+ RNA harvested from 8-day-old hydroponically grown, BTx623 sorghum seedlings. For some plants, one-half of the second leaf was crushed without damaging the midvein. For others, methyl jasmonate was added to the growth medium to a final concentration of 100 μ M. Leaves were harvested 3 and 27 hr after treatment and pooled. Double-stranded cDNA was cloned unidirectionally into different DraIII sites of the pME18S-FL3 vector (5-prime DraIII site is CACGTGTGG, 3-prime DraIII site is CACCATGTG). XhoI excises the cDNA insert."	

ORIGIN	Insert.."
Query Match	51.6%; Score 619.8; DB 7; Length 804;
Best Local Similarity	88.5%; Pred. No. 4.7e-139;
Matches 700; Conservative	0; Mismatches 77; Indels 14; Gaps 2;
Qy	412 ACTTGAACGGAGGTGTATTACTCTAGTAGCAATGATGGCTTCGAAAAAGTACTGTGG 471
Db	1 ACTTGAACGGAGGTGTATTACTCTAGTAGCAATGATGGCTTCGAAAAAGTACTGTGG 60
Qy	472 GGAAGATTATGTCGTAAGTCTCGGGTTATTCGTTCTTTTGATAGTGACAAAGTTAGTGGAGC 531

```

Db      61  GGAAGATTATGCTGAAGTCTTGGGTATTTCGTTCTTCACAGTGACAAATTAGTGGAGC 120
QY      532  AAGCTGTTGGAATGCCATCAGTTGGCCCAATATTCAGGTCATAGTGAAGCCTCTTTC 591
Db      121  AAGCTGTTGGAATGCCCTTCAGTTGCTCAATATATTCAGGTTTCATACGGAAGCCTCTTTC 180
QY      592  GGGAATAGAGAGTAGTGTCTTGAGAGATTGTCCTCCATCGACGACGATTAGTTGTGCCA 651
Db      181  GGGAATAGAGAGTAGTGTCTTGAGAGATTGTCCTCCATCGACGACGATTAGTTGTGCCA 240
QY      652  CCGAGGTGGTGTCTTTATCCGACCAATTAATCGGAGATATATGAAGAGGGGCTATCTG 711
Db      241  CCGCGGTGGTGTCTTTATCCGACCAATTAATCGGAAATATATGAAGAGGGGCTATCTG 300
QY      712  TTTGGTTAGATGTGCCCTTGGATGCTCTTGCTAGGCGTATTGCTAAAGTGGGAATGCCCT 771
Db      301  TTTGGTTAGTGTGCCCTTGGATGCTCTTGCTAGGCGTATTGCTAAAGTGGGAATGCCCT 360
QY      772  CTCGTCTCTCTTGGAACCAACCATCTGTGATCCGTAACGCAATGGCCTTTCTTAAGCTCA 831
Db      361  CTCGTCTCTCTTGGAACCAACCATCTGTGATCCATACACAAATGGCCTTTCTTAAGCTCA 420
QY      832  GCATGCTTCACAGCAAGAGGGTGATGCTTATGCAAAATGCAGATGAAGGTTTCTCTGG 891
Db      421  GCATGCTTCACAGCAAGAGGGTGACGCTTATGCAAAATGCAGATGAAGGTTTCTCTAG 480
QY      892  AAGAGATTGCATGTAACAAAGGTCATGATGATGCTCTTAAGCTGCACACCTACTGATATG 951
Db      481  AAGAGATTGCATGTAACAAAGGTCATGATGATGCTCTTAAGCTGCACACCTACTGATATG 540
QY      952  CAATTGAGTCACCTTCAATAGATCGAGAGCTTCTCATCGAGCACACTGCTGATAGTTCA 1011
Db      541  CAATTGAGTCACCTTCAATAGATCGAGAGCTTCTCATCGAGCACACTGCTGATATCCAG 600
QY      1012  CTAGCAGCGGCAAGCTGATGTCGAGATCCAGAGGATACAGACCTTGTGAACCTTAATC 1071
Db      601  CTAGCAGCTCGCAAGCTGATGTCGAGATCCAAAGGATACAGACCTTGTGAACCTTATC 660
QY      1072  CTTTCTTTGTTG-----CCACATAGACATCGTTGAGTTATTGCT--AAGGAT 1117
Db      661  CTTTCTTTGTTGACCGTAGTGTTACCCCTAGAGATCGTTGAGTTATTGCTCGTGCAT 720
QY      1118  GGAAGAGGAGGAGCTAATAATCCGAGTGTGCCCTCGCTCAAAAAAAGGAGGAGGAGG 1177
Db      721  GTCACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 780
QY      1178  AAAAAAAAAA 1188
Db      781  AGGGAGATATA 791

```

```

RESULT 3
CA078453
LOCUS
DEFINITION
  CA078453 667 bp mRNA linear EST 23-SEP-2003
  SCRLAM1007D04.g AM1 Saccharum officinarum cDNA clone SCRLAM1007D04
  5', mRNA sequence.
ACCESSION
  CA078453
VERSION
  CA078453.1 GI:34930725
SOURCE
  EST.
  Saccharum officinarum
  Saccharum officinarum
  Saccharum officinarum
  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
  clade; Panicoideae; Andropogoneae; Saccharum; Saccharum officinarum
  complex.
REFERENCE
  1 (bases 1 to 667)
  Vettore A.L., da Silva, F.R., Kemper, E.L. and Arruda, P.
  The libraries that made SUCEST
  Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
  Contact: Arruda P
  Centro de Biologia Molecular e Engenharia Genetica
  Universidade Estadual de Campinas

```

```

Caixa Postal 6010, 13083-970, Campinas SP, Brazil
Tel: 55 19 3788 1137
Fax: 55 19 3788 1089
Email: parruda@unicamp.br
Clone distribution: clone distribution information can be found
through the Brazilian Clone Collection Center (BCCC) at
http://www.bccccenter.fcav.unesp.br
Plate: 007 row: D column: 04
Seq primer: T7 Promoter Primer.
Location/Qualifiers
  1..667
  /organism="Saccharum officinarum"
  /mol_type="mRNA"
  /db_xref="taxon:4547"
  /clone="SCRLAM1007D04"
  /lab_host="DH10B"
  /clone_lib="AM1"
  /note="Organ: Apical meristem and tissues surrounding of
  mature plants; Vector: pSport1; Site_1: SalI; Site_2:
  NotI; An unidirectional cDNA library generated from
  [apical meristem and tissues surrounding of mature
  plants]. cDNA was prepared from polyA+ mRNA using
  SuperScript Plasmid System Kit (Invitrogen). The
  double-strand cDNAs were fractionated in a sepharose
  CL-2B 40cm-columns and fragments sizing between 0.8 and
  1.5 Kb were directionally cloned into the vector. Details
  of each source of RNA and library construction can be
  obtained at http://sucest.lad.ic.unicamp.br/public"

```

ORIGIN

```

Query Match      50.9%; Score 610.6; DB 6; Length 667;
Best Local Similarity 94.7%; Pred. No. 7.8e-137;
Matches 631; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

QY      340  GTCTGAAAACTCGCACAACTCCGTTGACGAAGCTCTCTGTTGAAGAGAAAAATCGAAG 399
Db      1  GTCTGAAAACTTCATTAACCTCCGTTGACGAAGCTCTCTTATTGAAGAGAAAAATCAGAAG 60
QY      400  AAGTTCTGTTCTACTTGAACGGGAGGTGTATTACTAGTAGGAATGATGGGTTCTGGAA 459
Db      61  AAGTTCTGTTCTACTTGAACGGGAGGTGTATTACTAGTAGGAATGATGGGTTCTGGAA 120
QY      460  AAGTACTGTGGGGAAGATTATGCTGAAGTCTCGGTTATTTCGTTTGTAGTAGTACA 519
Db      121  AAGTACTGTGGGGAAGATTATGCTGAAGTCTCGGTTATTTCGTTTGTAGTAGTACA 180
QY      520  AGTTAGTGAGCAAGCTGTTGGAATGCCATCAGTTGCCCAATATTAAGGTTCCATAGTG 579
Db      181  AATTAGTGAGCAAGCTGTTGGAATGCCCTTCAGTTGCCCAATATTAAGGTTCCATAGCG 240
QY      580  AAGCCTCTTTTCGGGATAATGAGAGTAGTGTCTTCGAGAGATTGTCTCCTCCATGCCAGCAT 639
Db      241  AAGCCTCTTTTCGGGATAATGAGAGTAGTGTCTTCGAGAGATTGTCTCCTCCATGCCAGCAT 300
QY      640  TAGTTGTTGCCACCGGAGGTGGTGTGTTATCCGACCAATTAACCTGGAGATATATGAAGA 699
Db      301  TAGTTGTTGCCACCGGAGGTGGTGTGTTATCCGACCAAGTTAACTGGAAATATATGAAGA 360
QY      700  GGGGCTATCTGTTTGGTTAGATGTCCTTGGATGCTCTGCTAGGCGGTATGCTAAG 759
Db      361  AGGGACTATCTGTTTGGTTAGATGTCCTTGGATGCTCTGCTAGGCGGTATGCTAAG 420
QY      760  TGGGAACCTGCTCTCGTCTCTTTCGGACCAACCATCTGGTATCCGTACGCAATGGCCCT 819
Db      421  TGGGAACCTGCTCTCGTCTCTTTCGGATCAACCATCTGGTATCCATACACAAATGGCCCT 480
QY      820  TTTCTAAGCTCAGCATGCTTGACAGCAAGAGGGGTGATGCTTATGCAAAATGCAGATGTA 879
Db      481  TTTCTAAGCTCAGCATGCTTGACAGCAAGAGGGGTGATGCTTATGCAAAATGCAGATGTA 540
QY      880  GGGTTTCTCTGGAAGAGATTGATGTAACAAAGGTCATGATGATGCTCTTAAGCTGACAC 939
Db      541  NGGTTTCTCTGGAAGAGATTGATGTAACAAAGGTCATGATGATGCTCTCTTAAGCTGACAC 600

```

QY 940 CTACTGATATGCAATTGAGTCACCTTCATAAGATCGAGAGCTTCGTCATCGAGCACACTG 999
 |||||
 Db 601 CCACCTGATATCGCAATTGAGTCACCTTCATAAGATCGAGAGCTTCGTCAGCGAGCACACTC 660
 |||||
 QY 1000 CTGATA 1005
 |||||
 Db 661 CAAATA 666

RESULT 4
 CAL146167
 LOCUS
 DEFINITION
 SCVPR2077B07.g RT2 Saccharum officinarum cDNA clone SCVPR2077B07
 5', mRNA sequence.

ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Saccharum officinarum
 Saccharum officinarum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoideae; Andropogoneae; Saccharum; Saccharum officinarum
 complex.

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT
 1 (bases 1 to 672)
 Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.
 The libraries that made SUCEST
 Genet. Mol. Biol. 24 (1-4), 1-7 (2001)

Contact: Arruda P
 Centro de Biologia Molecular e Engenharia Genetica
 Universidade Estadual de Campinas
 Caixa Postal 6010, 13083-970, Campinas SP, Brazil
 Tel: 55 19 3788 1137
 Fax: 55 19 3788 1089
 Email: parruda@unicamp.br
 Clone distribution: clone distribution information can be found
 through the Brazilian Clone Collection Center (BCCC) at
 http://www.bcccenter.fcav.unesp.br
 Plate: 077 row: E column: 07
 Seq primer: T7 Promoter Primer.

FEATURES
 source
 1..672
 /location/Qualifiers
 /organism="Saccharum officinarum"
 /mol_type="mRNA"
 /db_xref="taxon:4547"
 /clone="SCVPR2077B07"
 /lab_host="DH10B"
 /clone_lib="RT2"
 /note="Organ: Root tips(0.3cm-long) from adult plants;
 Vector: pSport1; Site 1: SalI; Site 2: NotI; An
 unidirectional cDNA library generated from [Root
 tips(0.3cm-long) from adult plants]. cDNA was prepared
 from polyA+ mRNA using SuperScript Plasmid System Kit
 (Invitrogen). The double-strand cDNAs were fractionated
 in a sepharose CL-2B 40cm-columns and fragments sizing
 between 0.8 and 1.5 Kb were directionally cloned into the
 vector. Details of each source of RNA and library
 construction can be obtained at
 http://sucest.lad.ic.unicamp.br/public"

ORIGIN
 Query Match 47.3%; Score 567.6; DB 6; Length 672;
 Best Local Similarity 93.5%; Pred. No 2e-126;
 Matches 604; Conservative 0; Mismatches 36; Indels 6; Gaps 1;

QY 130 CCGCGAGCGAACAATGAGCGGGGGCGTCGCTGCGCTGACGGCGCGGGCGCGG 189
 |||||
 Db 33 CCGCGAGCGACCAATGAGCGGGGGCGTGGCGCTGCGAGCGCGGGCGCGG 92
 |||||
 QY 190 GCTTCGGCTCCAGCGGACCGGGGGCGCTACAGCGGCCCGGAGCCTGAGAGTCG 249
 |||||
 Db 93 GCTTCGGCTCCAGCGGCGCGGGGGCGGCTACAGTCGCCGCGAGCCTGAGAGTCG 152
 |||||

QY 250 CTGACCCCGCGGACCTCGCGTCTGTCGGGGCTCGCGGGTCCAAAGCCCGTCGACCGC 309
 |||||
 Db 153 CTGGCCCGTCGGGAGCAGCGGTGCTGTGCGGGCTCGCGGGTCCAAAGCCCGTCGACCG- 211
 |||||
 QY 310 TCCGACTCCGTCGGAAGAAATCGTCCGAGGTCATGAAACTCGCACAACTCCGTTGAGC 369
 |||||
 Db 212 -----CTCCGTCGGAAGAAATCGTTCGGTGGTCATGAAAACTTGCATAACTCCGTTGAGC 266
 |||||
 QY 370 AAGCTCTCTCTGTTGAAGAGAAATCAGAAAGATGTTCTCTTCTTACTTGAACGGGAGGTGTA 429
 |||||
 Db 267 AAGCTCTCTCTTATTGAAGAGAAATCAGAAAGATGTTCTTCTTACTTTGAACGGGAGGTGTA 326
 |||||
 QY 430 TTTACTAGTAGGAATGATGGGTTCTCGAAAAGTACTGTGGGGAAGATTATGTCTGAAG 489
 |||||
 Db 327 TTTACTAGTAGGAATGATGGGTTCTCGAAAAGTACCGTGGGGAAGATTATGTCTGAAG 386
 |||||
 QY 490 TCTTGGGTATTTCGTTCTTTGATAGTAGTACAAAGTTAGTCGAGCAAGCTGTGGAATGCCAT 549
 |||||
 Db 387 TCTTGGGTATTTCGTTCTTCGATAGTGACAAATTAGTCGAGCAAGCTGTGGAATGCCCTT 446
 |||||
 QY 550 CAGTTGCCCCAATATTCAAGGTCCATAGTGAAGCCCTTCTTCGGGATATAGAGAGTAGTG 609
 |||||
 Db 447 CAGTTGCCCAATATTCAAGGTTCATAGCGAAGCCCTTCTTCGGGATATAGAGAGTAGTG 506
 |||||
 QY 610 TCTTGAGAGATTGTCTCCATGCGACGATTAGTTGTTGCCACCGAGGTGGTGTCTGTTA 669
 |||||
 Db 507 TCTTGAGAGATTGTCTCCATGCGACGATTAGTGGTTGCCACCGAGGTGGTGTCTGTTA 566
 |||||
 QY 670 TCCGACCAATTAACCTGGAGATATAGAAAGGGGCTATCTGTTTGGTTAGATGTCCTT 729
 |||||
 Db 567 TCCGACCAAGTTAACTGGGAATATATGAAGAAGGACTATCTGTTTGGTTAGATGTCCTT 626
 |||||
 QY 730 TGGATGCTCTTCTAGGCGTATTGCTAAAGTGGGAACCTGCCTCTCG 775
 |||||
 Db 627 TGGATGCTCTGCTAAGCGTATTGCTAAAGTGGGAACCTGCCTCTCG 672
 |||||

RESULT 5
 BQ577650
 LOCUS
 DEFINITION
 lab Zea mays cDNA, mRNA sequence.
 3524.1_39.1_F07.Y.1_3524 - Mature pollen from Sheila McCormick's

ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 BQ577650.1 GI:21480967
 BQ577650
 EST.
 Zea mays
 Zea mays
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoideae; Andropogoneae; Zea.

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT
 1 (bases 1 to 582)
 Walbot,V.
 Maize ESTs from various cDNA libraries sequenced at Stanford
 University
 Unpublished (1999)
 Contact: Walbot V
 Department of Biological Sciences
 Stanford University
 855 California Ave, Palo Alto, CA 94304, USA
 Tel: 650 725 8221
 Fax: 650 725 8221
 Email: walbot@stanford.edu
 Plate: 3524.1_39.1 row: F column: 07.

FEATURES
 source
 1..582
 /organism="Zea mays"
 /mol_type="mRNA"
 /cultivar="B73"
 /db_xref="taxon:4577"
 /tissue_type="pollen"
 /dev_stages="mature"
 /lab_host="SOLR"
 /clone_lib="3524 - Mature pollen from Sheila McCormick's

lab"		/notes="Vector: Stratagene's Uni-Zap XR (pBluescript SK-); Site 1: EcoRI (5-prime); Site 2: XhoI (3-prime); Unamplified cDNA library directionally cloned by Rima Kulikauskas using Stratagene's Uni-Zap system. Insert sizes ranged from 0.5Kb to 2Kb. 50 microliter aliquot had 330,000 pfu when it was made in Sept, 1995, from oligo dn-primed poly A+ RNA."	
ORIGIN		Query Match 47.2%; Score 566; DB 5; Length 582; Best Local Similarity 100.0%; Pred. No. 4.8e-126; Matches 566; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1	CCGCCACCACTACCTGCTCTCTCTCTCTCTCTCTTTTACACCTCACCTCCGCGATCGCTC	60
Db	17	CGGCCACCACTACCTGCTCTCTCTCTCTCTCTCTTTTACACCTCACCTCCGCGATCGCTC	76
QY	61	AGAGAGTCAGAGATTCGAGTTGAGCTATAGGCGTAGCCGAGCTGGTCCGCGGTCCCTCT	120
Db	77	AGAGAGTCAGAGATTCGAGTTGAGCTATAGGCGTAGCCGAGCTGGTCCGCGGTCCCTCT	136
QY	121	CGCTCCACCGCGGAGCGAACAATGAGCGCGGGGCGCTCGGCTGGCGCTGCAGGCGC	180
Db	137	CGCTCCACCGCGGAGCGAACAATGAGCGCGGGGCGCTCGGCTGGCGCTGCAGGCGC	196
QY	181	GGCGCGGCGCTTCGGCTCCAGCGCGCACCGGGCGGCTTACAGGCGCCACCGGGAGCC	240
Db	197	GGCGCGGCGCTTCGGCTCCAGCGCGCACCGGGCGGCTTACAGGCGCCACCGGGAGCC	256
QY	241	TGAGAGTCGTGACCGCGGGGACCTGCGGTGCTGTGCGGGTCCGCGGTCAAGCCCG	300
Db	257	TGAGAGTCGTGACCGCGGGGACCTGCGGTGCTGTGCGGGTCCGCGGTCAAGCCCG	316
QY	301	TGCGACCGCTCGGACTCCGTGCGAAGAAATCGTCCGAGGTCAATGAAGCTCGCAACT	360
Db	317	TGCGACCGCTCGGACTCCGTGCGAAGAAATCGTCCGAGGTCAATGAAGCTCGCAACT	376
QY	361	CGCTTGACGAAGCTCTCTCTTTGAAGAGAAATCAGAAGAGTTCTGTCTACTTTGAACG	420
Db	377	CGCTTGACGAAGCTCTCTCTTTGAAGAGAAATCAGAAGAGTTCTGTCTACTTTGAACG	436
QY	421	GGAGGTGATTTACCTAGTAGAATGATGGTTCTGGAAAAAGTACTGTGGGAAGATTA	480
Db	437	GGAGGTGATTTACCTAGTAGAATGATGGTTCTGGAAAAAGTACTGTGGGAAGATTA	496
QY	481	TGCTCAAGCTCTGGGTATTCTGTTCTTGATAGTACAAAGTTAGTGGAGCAAGCTGTTG	540
Db	497	TGCTCAAGCTCTGGGTATTCTGTTCTTGATAGTACAAAGTTAGTGGAGCAAGCTGTTG	556
QY	541	GAATGCCATCAGTTGCCCAATATTC	566
Db	557	GAATGCCATCAGTTGCCCAATATTC	582
RESULT 6		CB684025	
LOCUS		OSJNEf13C12.f OSJNEf Oryza sativa (japonica cultivar-group) cDNA	
DEFINITION		clone OSJNEf13C12 5', mRNA sequence.	
ACCESSION		CB684025	
VERSION		1	
KEYWORDS		EST.	
SOURCE		Oryza sativa (japonica cultivar-group)	
ORGANISM		Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.	
REFERENCE		1 (bases 1 to 825)	
AUTHORS		Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E., Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.	
TITLE		Large-scale identification of ESTs involved in the interaction between rice and Magnaporthe grisea	
JOURNAL COMMENT		Unpublished (2003) Contact: Rod Wing Arizona Genomics Institute University of Arizona Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ 85721-0088, USA Tel: 520 626 3967 Fax: 520 621 9288 Email: http://genome.arizona.edu PCR Primers FORWARD: gta aaa cga cgg cca gtcg BACKWARD: gga aac agc tat gac cat g Plate: 13 row: C column: 12 Seq primer: gta aaa cga cgg cca gtcg. Location/Qualifiers 1. 825 /organism="Oryza sativa (japonica cultivar-group)" /mol_type="mRNA" /cultivar="Nipponbare" /db_xref="taxon:39947" /clone="OSJNEf13C12" /tissue_type="Leaf" /dev_stage="3 week" /lab_host="DH10B" /clone_lib="OSJNEf" /note="Vector: pBluescript II KS +; Site_1: EcoRI; Site_2: XhoI; Uninfected Control"	
FEATURES		source	
ORIGIN		Query Match 47.2%; Score 566; DB 6; Length 825; Best Local Similarity 84.3%; Pred. No. 5e-126; Matches 670; Conservative 0; Mismatches 101; Indels 24; Gaps 2;	
QY	149	GGCGGGGGCGCTGCGGCTGCGGCTGCGGCGGGCGGGCGGCTTCGGCTCCAGCGCGCA	208
Db	36	GGAGCGGGCGGTGGGGCTGCGCTGAGTCCGGCGGGCGGGTTCGGCGGTCCGACCG	95
QY	209	CGCG-----GGCGGCTACAGCGGCCACCGGGAGCTCGAGAGTCGC	250
Db	96	CGCGCGGAGCGGCTCTACGGCGGCGAGGGCGGGCGGAGCTTCGAGGTGCG	155
QY	251	TCACCGGGGGAGCTGCGGTGCTGTGGGGCTCGGGGTCAAAGCCGTCGACCGCT	310
Db	156	TGAGCGCGGGTGGGAAAGCGCGCTGTGTGGGCTCGCGGTCCAAAGCGGTCGCCCG	213
QY	311	CGGACTCGTGGAGAAATCGTCCGAGGTCAATAAACTCGCAACTCGCTTGACCA	370
Db	214	-----CTCCGTGCCAAGAAATCGTCGGAGGTCAATAAACTCGCTTGATGA	269
QY	371	AGCTCTCTCTTTGAAGAGAAATCAGAAGAGTTCTGTCTTACTTTGAACGGGAGGTGTAT	430
Db	270	AGCCCTCTTCTAAAGAGAAATCAGAAGAGTTCTCTTCTATTGTAATGGACGGTGTAT	329
QY	431	TTACCTAGTAGAATGATGGGTTCTGGAAAAAGTACTGTGGGGAAGATTAATCTGAACT	490
Db	330	TTACCTAGTTGAATGATGGGTTCTGGAAAAAGTACTGTGGGGAAGATCATGCTGAACT	389
QY	491	CTTGGGTTATGCTTCTTTGATAGTACAAAGTTAGTGGAGCAAGCTTGTGGAATGCCATC	550
Db	390	TTTGGGTTATGCTTCTTTGATAGTAAATTTGTCGAACCAAGCTGTGGGATGCTCTTC	449
QY	551	AGTTCGCCAAATATTCAAGGTCATAGTGAAGCCCTTTCTTCGGGATAATGAGAGTAGTGT	610
Db	450	AGTCCTCAAAATTTTCAAGGTTCAATAGTGAAGCCCTTTTAGGGATAATGAGAGTAGTGT	509
QY	611	CTTGAGAGATTGCTCTCCATCGCAGATAGTGTGTCACCGGAGGTGGTGTGTAT	670
Db	510	CTTGAGGAGTTGCTCTCAATGAAGCGATAGTGTGCTGCTACTGAGAGGTGGTGTGTAT	569
QY	671	CGGACCAATTAACTCGAGATATATCAAGGGGCGCTATCTGTTCTGTTAGATGTCCTT	730
Db	570	CGGACCAATTAACTCGGAATACATGAAGAGGGCCCTATCTGTTGTTGGATGTCCTT	629

QY 731 GGATGCTCTTGTAGCGGATTAGTAAAGTGGAACTGCCTCTCGTCTCTTCTGACCA 790
 Db 630 GGACGCTCTTGTAGCGGATTAGTCTANAGTGGGACTGCCTCCGTCCTTCTTAGATCA 689
 QY 791 ACCATCTGGTATCCGTACGCAATGGCCCTTTTCTAAGCTCAGCATGCTTGCACAGCAAG 850
 Db 690 ACCATCTGGTATCCATACACAAATGGCTTTTCTAAGCTCAGCATGCTCCTCGGAGCAAG 749
 QY 851 GGGTATGCTTATGCAAAATGCAAGATGAAGGTTTCTCTGGAAGAGATTGCATGTAAACA 910
 Db 750 GGGCCATGCTTATGCAAAATGCTGATGTAGGGTTTCTCTTGAAGAGATTGCATCTAAACA 809
 QY 911 AGGTCAATGATGT 925
 Db 810 GGGTCATGATGATGT 824

RESULT 7
 COS24949
 LOCUS
 DEFINITION
 3530.1_165.1_C06.Y.1 3530 - Full length cDNA library created by
 Invitrogen from multiple tissues Zea mays cDNA, mRNA sequence.
 COS24949
 COS24949.1 GI:50329823

ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Zea mays
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoideae; Andropogoneae; Zea.
 1 (bases 1 to 549)

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT
 Walbot,V.
 Maize ESTs from various cDNA libraries sequenced at Stanford
 University
 Unpublished (1999)
 Contact: Walbot V
 Department of Biological Sciences
 Stanford University
 855 California Ave, Palo Alto, CA 94304, USA
 Tel: 650 723 2227
 Fax: 650 725 8221
 Email: walbot@stanford.edu
 Plate: 3530.1.165.1 row: C column: 06.

FEATURES
 source
 1..549
 /organism="Zea mays"
 /mol_type="mRNA"
 /cultivar="B73"
 /db_xref="taxon:4577"
 /tissue_type="multiple"
 /dev_stage="varies by tissue"
 /lab_host="DH10B"
 /clone_lib="3530 - Full length cDNA library created by
 Invitrogen from multiple tissues"
 /note="Organ: silks, husks, ears, pollen, shoot tips,
 leaf, root tips, whole seed, embryo; Vector: pCMV-SPORT
 6.1; Site_1: EcoRV; Site_2: NotI; Maize Gene Discovery
 Project contracted with Invitrogen to produce a
 normalized, full length library in a pSport vector. This
 is a Gateway compatible vector, permitting clone movement
 to new vector backbones for expression in diverse host
 cells using recombination rather than restriction enzymes.
 Details of the vector and sequencing primers are available
 at ZmDB in the EST library description tables. poly(A)+
 mRNA was prepared by Invitrogen, and equimolar amounts of
 RNA from each of the 12 tissue samples were mixed together
 for selection of mRNA with a 5' cap. After synthesis of
 cDNA, a normalization step was conducted against the
 mixture of RNA sources. This step effected a 20X to 80X
 reduction in common transcript types. Tissues prepared: 1.
 just emerging silks; 2. inner husks from ears of sample
 #1; 3. 20 dap aleurone; 4. immature tassels, stages from
 1-2 mm to 1-2 cm; 5. 2 mm to 2 cm ears; 6. pollen; 7. 1 cm

ORIGIN

Query Match 45.8%; Score 549; DB 7; Length 549;
 Best Local Similarity 100.0%; Pred. No. 6.2e-122;
 Matches 549; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 98 CGACTGCTCGCGCGTCCCTCTCGCTCCACCCGGCGGCAACAATGGAGCGGGGG 157
 Db 1 CGACTGGTGGCGCGTCCCTCTCGCTCCACCCGGCGGCAACAATGGAGCGGGGG 60
 QY 158 CGTGGCCCTGGCGCTGCAGGGCGCGGGGCTTCGGCTCCAGCCCGGCACCGGGGG 217
 Db 61 CGTGGCCCTGGCGCTGCAGGGCGCGGGGCTTCGGCTCCAGCCCGGCACCGGGGG 120
 QY 218 CTTACAGGGCCCAACCGGAGCTGAGAGTGCCTGACCCGGCGGACCTCGGTCTGT 277
 Db 121 CTTACAGGGCCCAACCGGAGCTGAGAGTGCCTGACCCGGCGGACCTCGGTCTGT 180
 QY 278 GCGGCTCGGGTCCAAAGCCGTCGACCCGCTCCGCTCCGTCGGAAGAAATCTCCGG 337
 Db 181 GCGGCTCGGGTCCAAAGCCGTCGACCCGCTCCGCTCCGTCGGAAGAAATCTCCGG 240
 QY 338 AGGTTCATGAAAACTCGCAACTTCGGTTGACGAAGCTCTCTCTGTTGAAGAGAAAAATCAGA 397
 Db 241 AGGTTCATGAAAACTCGCAACTTCGGTTGACGAAGCTCTCTCTGTTGAAGAGAAAAATCAGA 300
 QY 398 AGAAGTCTCTTCTACTCTTGAACGGGAGGTGATTTACCTAGTAGGAATGAGGTTCTGG 457
 Db 301 AGAAGTCTCTTCTACTCTTGAACGGGAGGTGATTTACCTAGTAGGAATGAGGTTCTGG 360
 QY 458 AAAAGTAGTCTGGGGAAGATTATCTGCAAGCTTTGGTTATTCTGTTCTTTGATAGCA 517
 Db 361 AAAAGTAGTCTGGGGAAGATTATCTGCAAGCTTTGGTTATTCTGTTCTTTGATAGCA 420
 QY 518 CAAAGTTAGTGAGCAAGCTGTTTGGAAATGCCATCAGTTGCCAAATATTCAAGGTCCATAG 577
 Db 421 CAAAGTTAGTGAGCAAGCTGTTTGGAAATGCCATCAGTTGCCAAATATTCAAGGTCCATAG 480
 QY 578 TGAAGCCTTCTTTTCGGGATATGAGAGTAGTGTCTTTGAGAGATTGTCTCTCATGCGAGC 637
 Db 481 TGAAGCCTTCTTTTCGGGATATGAGAGTAGTGTCTTTGAGAGATTGTCTCTCATGCGAGC 540
 QY 638 ATTAGTTGT 646
 Db 541 ATTAGTTGT 549

RESULT 8

CA452817/c
 LOCUS
 DEFINITION
 Rp3A-3_H01 subtracted cDNA library of maize inbred line H95-Rp3A
 inoculated with Puccinia sorghi isolate IN1 Zea mays cDNA clone
 Rp3A-3_H01, mRNA sequence.
 CA452817
 CA452817.1 GI:24934599
 EST.
 Zea mays
 Zea mays
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Eukaryota; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD

vegetative shoot tips from 15 day old seedlings; all
 leaves with an expanded or partially expanded sheath
 were removed; 8. mature leaf tissue; 9. 0.5 cm long root
 tips from 15 day old seedlings; 10. 10 dap whole seed; 11.
 12 dap endosperm and embryo; 12. 17 dap endosperm and
 embryo. All of the sequenced clones in project 3530 will
 be archived at the University of Arizona along with the
 Unigene clones from the Maize Gene Discovery EST
 sequencing projects. Clones can be ordered through the
 ZmDB web site or directly from the University of Arizona
 (http://www.genome.arizona.edu/orders/). High density
 filters containing over 18,000 clones can also be ordered
 from the University of Arizona."

clade; Panicoidae; Andropogoneae; Zea.
1 (bases 1 to 619)
Bai, J., Fellers, J.P., Leach, J.E. and Hulbert, S.H.
Comparison of pathogen induced defense gene profiles on maize lines
with different resistance genes
Unpublished (2003)
JOURNAL
COMMENT

Contact: Bai J
Department of Plant Pathology
Kansas State University
4024 Throckmorton Plant Sciences Center, Manhattan, KS 66506, USA
Tel: 785-532-2328
Fax: 785-532-5692
Email: jianfa@plantpath.ksu.edu
Seq primer: T7.

FEATURES
source
1..619
Location/Qualifiers
/organism="Zea mays"
/mol_type="mRNA"
/strain="inbred line H95-Rp3A"
/db_xref="taxon:4577"
/clone="Rp3A-3 H01"
/tissue_type="Rust-infected leaves"
/dev_stage="4 week-old plants"
/clone_lib="subtracted cDNA library of maize inbred line
H95-Rp3A inoculated with Puccinia sorghi isolate IN1"
/notes="Vector: pUC19; Suppression subtractive
hybridization; cloned into pUC19 vector. From a subtracted
cDNA library of maize inbred line H95 carrying the Rp3A
gene inoculated with an incompatible Puccinia sorghi
isolate."

ORIGIN

Query Match 45.6%; Score 547.8; DB 6; Length 619;
Best Local Similarity 93.2%; Pred. No. 1.2e-121;
Matches 573; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY 465 ACTGTGGGGAAGATTATGCTGAAGCTTTGGGTTATTGCTTTGATAGTACAAAGTTA 524
DB 619 ACTGTGGGGAAGATCATGCTGAAGCTTTGGGTTATTGCTTTGATAGTACAAATTA 560
QY 525 GTGAGCAGCTTTGGATGCCATCAGTTGCCAATATTCAAGTCCATAGTGAAGCC 584
DB 559 GTGAGCAGCTTTGGATGCCCTCAGTTGCTCAAAATATTCAAGTTACAGTGAAGCC 500
QY 585 TTCTTTCCGGATAATGAGAGTAGTGTCTTGAGAGATTGTGCTCCATGCGACGATTAGTT 644
DB 499 TTCTTTCCGGATAATGAGAGTAGTGTCTTGAGAGATTGTGCTCCATGCGACGATTAGTT 440
QY 645 GTTGCCACCGAGGTGGTGTCTTATCCGACCAATTAACCTGGAGATATATGAAGAGGGC 704
DB 439 GTTGCCACCGAGGTGGTGTCTGTCATCCGACCAAGTTAACTGGAAATATATGAAGAGGGC 380
QY 705 CTATCTGTTGGTTAGATGTCCTTGATGCTCTTGCTAGGCTATTCGTAAGTGGGA 764
DB 379 CTATCGGTTTGGTTAGATGTCCTTGATGCTCTTGCTAGGCTATTCGTAAGTGGGA 320
QY 765 ACTGCTCTGCTCTCTTCTGGACCAACCATCTGCTGATCCGACCAATGCGCTTTTCT 824
DB 319 ACCGCTCTGCTCTCTTCTGGACCAACCGCTCGGTGATCCATACAAATGCGCTTTTCT 260
QY 825 AAGCTCAGCATGCTTGACAGCAAAAGGGGTGATGCTTATGCAAAATCAGATGTAAGGGTT 884
DB 259 AAGCTCAGCATGCTTGACAGCAAAAGGGGTGATGCTTATGCAAAATCGGATGTAAGGGTT 200
QY 885 TCTCTGAGAGATTGATGTAACAAAGGTTCATGATGATGTCTCTAAGCTGACACTACT 944
DB 199 TCTCTGAGAGATTGATGTAACAAAGGTTCATGATGATGTCTCTAAGCTGATGCGGACT 140
QY 945 GATATTGCAATTCAGTCACTTCAATAGATCGAGAGCTTCGTCTATCGAGCACACTGCTGAT 1004
DB 139 GATATTGCAATTCAGTCACTTCAATAGATCGAGAGTTTCGTCTATCGAGCACGCTGCTGAT 80
QY 1005 AGTTTCACTAGCGACGCGCAAGCTGAGTCCAGATCCAGAGGATACAGACCTTGTAGAAC 1064

DB 79 AATCCAGCTAGGCACTCGCAAGCTAGTCACAGATCCAAAGGATACAGACCTTGTAAATAT 20
QY 1065 CTTAATCCCTTTGTT 1079
DB 19 CTTAATCCCTTTGTT 5

RESULT 9
CB657302
LOCUS
DEFINITION
OSJNEC12117.f OSJNEC Oryza sativa (japonica cultivar-group) cDNA
clone OSJNEC12117 5', mRNA sequence.
ACCESSION
CB657302
VERSION
CB657302.1 GI:29661027
KEYWORDS
EST.
SOURCE
Oryza sativa (japonica cultivar-group)
ORGANISM
Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Erbartoideae; Oryzeae; Oryza.
1 (bases 1 to 833)
REFERENCE
AUTHORS
Jantassuriyarat, C., Lu, G., Gowda, M., Hatfield, J., Zhou, B., Mazur, E.,
Kudrna, D., Dean, R., Soderlund, C., Wing, R. and Wang, G.
TITLE
Large-scale identification of ESTs involved in the interaction
between rice and Magnaporthe grisea
JOURNAL
Unpublished (2003)
COMMENT
Contact: Rod Wing
Arizona Genomics Institute
University of Arizona
Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
85721-0088, USA
Tel: 520 626 3967
Fax: 520 621 9288
Email: http://genome.arizona.edu
PCR Primers
FORWARD: gta aaa cga cgg cca gtc
BACKWARD: gga aac agc tat gac cat g
Plate: 12 row: I column: 17
Seq primer: gta aaa cga cgg cca gtc.

FEATURES
source
1..833
Location/Qualifiers
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="OSJNEC12117"
/tissue_type="Leaf"
/dev_stage="3 week"
/lab_host="DH10B"
/clone_lib="OSJNEC"
/note="Vector: pBluescript II KS +; Site 1: EcoRI; Site 2:
XhoI; 6 hrs after inoculation with Rice Blasc (C9240-1)"

ORIGIN

Query Match 45.6%; Score 546.6; DB 6; Length 833;
Best Local Similarity 84.1%; Pred. No. 2.5e-121;
Matches 650; Conservative 0; Mismatches 99; Indels 24; Gaps 2;

QY 149 GCGCGGGGCGTTCGCGCTGCGCTCAGCGCGGGCGGCGGCTTCGGTCCAGCGGCA 208
DB 67 GAGAGCGGCGTGGGCTGGCGCTGCGCTCAGTCGCGGGCGGCGGCTTCGGCGCTCCGACCG 126
QY 209 CCGG-----GGCGGCTACAGGGCCCAAGGAGCTCAGAGTCGC 250
DB 127 CCGCGGAGCGCGCTTACGCGCGGCGAGGGCGGCGGATCGGGAGCTTGAAGGCTCGC 186
QY 251 TGACCGCGGCGGACTTCGCGCTGCTGCGGCTCGCGGTCGCGGTCGCGGTCGCGGCTCGCGGCT 310
DB 187 TGAGCGGCGGTTGGCGAGGCGGCTGTTGGGCTCGGGTCAAGCGGTCGCGGCTCGCGGCT 244
QY 311 CCGACTCGTGCAGAAATTCGTCGGAGGTTCATGAAACTCGCAACTCGCGTTGACGA 370

Db 245 ----CTCCGTGCAAGAAATCGTCCGAGGTGATGAACATTCGATCAATGCAATCGGTTGATGA 300

Qy 371 AGCTCTCTGTTGAAGAGAAAATCAGAAAGTTCTGTTCTACTTCTACTTGAACGGAGGTGAT 430

Db 301 AGCCCTCTTGTCTAAAGAGAAAATCAGAAAGTTCTTCTTATTTGAATGGAGGTGAT 360

Qy 431 TTACCTTAGTAGGAATGATCGGTTCTGGAAGAGTACTGTTGGGAAGATTATGCTGAAAGT 490

Db 361 TTACCTTAGTAGGAATGATCGGTTCTGGAAGAGTACTGTTGGGAAGATTATGCTGAAAGT 420

Qy 491 CTTGGGTTATTCGTTCTTTGATGATGACAAAGTTAGTGGAGCAAGCTGTTGGAATGCCATC 550

Db 421 TTTGGGTTATTCGTTCTTTGATGATGATAAATTTGGTTCGAACAAGCTGTGGCATGCTTC 480

Qy 551 AGTTGCCCAAAATATCAAGGTCATAGTGAAGGCTTCTTTCGGGATAATGAGAGTAGTGT 610

Db 481 AGTCGCTCAAAATTTTCAAGGTTTCATAGTGAAGGCTTCTTTAGGGATAATGAGAGTAGTGT 540

Qy 611 CTTGAGAGATTGTCCTCATCGACGATTTAGTTGTTGTCACCGGAGGTGGTCTGTTAT 670

Db 541 CTTGAGGATTGTCCTCAATGAAGGATTTAGTTGTTGCTACTGAGGTTGGTCTGTTAT 600

Qy 671 CCGACCAATTAATCGGAGATATATGAAGGGGCCCTATCTGTTGGTTAGATGTCGCCCT 730

Db 601 CCGACCAATTAATCGGAGATATATGAAGGGGCCCTATCTGTTGGTTAGATGTCGCCCT 660

Qy 731 GGATGCTTCTGAGCGTATGCTTAAAGTGGAACTGCTCTGCTCTCTTCTGGAACA 790

Db 661 GGACGCTCTTGTAGCGTATTTGCTAAAGTGGGAGTGCCTCCGCTCTCTTCTAGATCA 720

Qy 791 ACCATCTGTGTATCGCTGCAATGCGCTTTTCTAAGCTCAGATGCTTGCACAGCAAG 850

Db 721 ACCATCTGTGTATCGCTGCAATGCGCTTTTCTAAGCTCAGATGCTTGCACAGCAAG 780

Qy 851 GGGTGATGCTTATGCAATGCAATGATGAAGGTTTCTCTGGAAGAGATTGCAT 903

Db 781 GGGCGATGCTTATGCAATGCTGATGCTGAAGGTTTCTCTTGAAGAGATTGCAT 833

RESULT 10

AW671996

LOCUS LG1.353 All.b1_A002 Light Grown 1 (LGI) Sorghum bicolor cDNA, mRNA

DEFINITION

ACCESSION AW671996.1 GI:7535901

VERSION

KEYWORDS

SOURCE Sorghum bicolor (sorghum)

ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Sorghum.

1 (bases 1 to 609)

Cordonnier-Pratt M.-M., Gingle, A., Marsala, C. and Pratt, L.H.

An EST database from Sorghum: light-grown seedlings

Unpublished (2000)

Contact: Cordonnier-Pratt MM

Laboratory for Genomics and Bioinformatics

The University of Georgia, Department of Plant Biology

Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA

Tel: 706 542 1860

Fax: 706 583 0210

Email: mpratt@uga.edu

Sequences have been trimmed to exclude PolyA, vector and regions below Phred quality 16. The threshold for highest quality sequence is 20.

Seq primer: JEN REV

High quality sequence stop: 607

POLYA=No. Location/Qualifiers

1..609

/organism="Sorghum bicolor"

/mol_type="mRNA"

FEATURES

source

1045..1099

ORIGIN

Query Match 45.2%; Score 542; DB 2; Length 609;

Best Local Similarity 94.1%; Pred. No. 3.1e-120;

Matches 576; Conservative 0; Mismatches 30; Indels 6; Gaps 1;

Qy 259 CGGACCTCGCGTCTGCGGGCTCGCGGGTCCAGCGCTCCACCGCTCCGACTCC 318

Db 4 CGAGGGCTGCGGTTGCCGTGCGGGTCCAGCGCTCCACCG-----CTCC 57

Qy 319 GTGCCAAGAAATCGTCCGAGGTCAATGAAACTCGCACAACTCCGTTGACGAAGCTCTCC 378

Db 58 GTGCCAAGAAATCGTTCGAGGTCAATGAAACTTCGATAACTCCGTTGACGAAGCTCTCC 117

Qy 379 TGTGAAGAGAAAATCAGAAAGTTCTGTTCTACTTGAACGGAGGTGTTATTTACCTAG 438

Db 118 TGTGAAGAGAAAATCAGAAAGTTCTGTTCTACTTGAACGGAGGTGTTATTTACTTAG 177

Qy 439 TAGGAATCATGGTTCTGGAAGAACTACTGTGGGGAAGATTATGCTGAAGTCTTTGGGT 498

Db 178 TAGGAATCATGGTTCTGGAAGAAAGTACAGTGGGGAAGATTATGCTGAAGTCTTTGGGT 237

Qy 499 ATTCGTTCTTTGATAGTAGTGAAGTTAGTGGAGCAAGCTGTTGGAATGCCATCAGTTGCC 558

Db 238 ATTCGTTCTTTGACAGTGAAGTTAGTGGAGCAAGCTGTTGGAATGCCATCAGTTGCC 297

Qy 559 AAATATTCAGGTTCATAGTGAAGCTTCTTTCCGGATAATGAGAGTAGTGTCTTGAGAG 618

Db 298 AAATATTCAGGTTCATAGCGAAGCTTCTTTCCGGATAATGAGAGTAGTGTCTTGAGAG 357

Qy 619 ATTTGCTCTCCATCGACGATTAGTTGTCACCGGAGGTGGTCTGTTATCCGACCAA 678

Db 358 ATTTGCTCTCCATCGACGATTAGTTGTCACCGGAGGTGGTCTGTTATCCGACCAA 417

Qy 679 TTAAGTGGAGATATATGAAGGGGCCCTATCTGTTGGTTAGATGTCCTTTGGATGCTC 738

Db 418 TTAAGTGGAGATATATGAAGGGGCCCTATCTGTTGGTTAGATGTCCTTTGGATGCTC 477

Qy 739 TTGCTAGGCGTATTCGTTAAAGTGGAACTGCTCTCGTCTCTTTTGACCAACCATCTG 798

Db 478 TTGCTAGGCGTATTCGTTAAAGTGGAACTGCTCTCGTCTCTTTTGACCAACCATCTG 537

Qy 799 GTGATCCGTACGCAATGCGCTTTCTAAGCTCAGCATGCTTGCACAGCAAGGGGTGATG 858

Db 538 GTGATCCGTACGCAATGCGCTTTCTAAGCTCAGCATGCTTGCACAGCAAGGGGTGATG 597

Qy 859 CTTATGCAAAATG 870

Db 598 CTTATGCAAAATG 609

RESULT 11

CF033002/c

LOCUS CF033002

DEFINITION

ACCESSION CF033002

VERSION

KEYWORDS

SOURCE

ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.

1 (bases 1 to 620)

Genoplante, a major partnership french program in plant genomics

/db_xref="taxon:4558"

/clone_lib="Light Grown 1 (LGI)"

/note="Organ: 10- to 14-day-old light-grown (greenhouse) seedlings; Vector: Lambda Zap; Site1: XhoI; Site2: EcoRI; The library was made from poly-A RNA in the cloning vector lambda ZAP II. Clones to be sequenced were prepared by mass excision."

JOURNAL COMMENT	Unpublished (2003) Contact: Genoplatte Genoplatte 93, rue Henri Rochefort 91025 EVRY CEDEX France Tel: 33 1 69 47 54 00 Fax: 33 1 69 47 54 10 This sequence has been generated in the framework of the french plant genomics programme 'Genoplatte' (http://www.genoplatte.com) and http://genoplatte-info.infobiogen.fr/ .		ACCESSION VERSION KEYWORDS SOURCE ORGANISM	CF636791.1 CF636791.1 EST. Zea mays Zea mays Zea mays	
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.				
FEATURES source	Location/Qualifiers 1..620 /organism="Zea mays" /mol_type="mRNA" /cultivar="F2" /db_xref="taxon:4577" /clone="QCF11b08" /tissue type="seedling minus kernel" /clone_lib="QCF"		REFERENCE AUTHORS	1 (bases 1 to 814) Bohnert, H., Sharp, R.E., Springer, G.K., Poroyko, V., Fredricksen, M., Sharp, L.G., Spollen, W.G., Ries, J., Guillen, A., Khambati, A., Topinka, C., Davis, G.E., Schachtman, D., Wu, Y. and Nguyen, H.T. NSF Grant DBI-0211842: Functional Genomics of Root Growth and Root Signaling Under Drought Unpublished (2003) Contact: Hans Bohnert University of Illinois, Urbana-Champaign 1201 West Gregory Drive, Urbana, IL 61801, USA Tel: 217-265-5475 Fax: 217-333-5574 Email: bohnert@life.uiuc.edu POLYA=Yes.	
	ORIGIN				
Query Match 45.1%; Score 540.8; DB 6; Length 620; Best Local Similarity 92.4%; Pred. No. 6e-120; Matches 569; Conservative 0; Mismatches 47; Indels 0; Gaps 0;					
QY	465 ACTGTGGGGAAGATTATGCTGAAGTCTTGGGTATTTCGTTCTTTGATAGTGAAGTTA 524		Location/Qualifiers		CF636791.1 GI:37398941 Zea mays Zea mays Zea mays Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea. 1 (bases 1 to 814) Bohnert, H., Sharp, R.E., Springer, G.K., Poroyko, V., Fredricksen, M., Sharp, L.G., Spollen, W.G., Ries, J., Guillen, A., Khambati, A., Topinka, C., Davis, G.E., Schachtman, D., Wu, Y. and Nguyen, H.T. NSF Grant DBI-0211842: Functional Genomics of Root Growth and Root Signaling Under Drought Unpublished (2003) Contact: Hans Bohnert University of Illinois, Urbana-Champaign 1201 West Gregory Drive, Urbana, IL 61801, USA Tel: 217-265-5475 Fax: 217-333-5574 Email: bohnert@life.uiuc.edu POLYA=Yes.
Db	620 ACTGTGGGAGATCCATGTCTGAAGTCTTGGGTATTTCGTTCTTTGATAGTGAAGTTA 561		1..814 /organism="Zea mays" /mol_type="mRNA" /db_xref="taxon:4577" /clone_lib="zmrw00" /note="Samples were collected in Robert E. Sharp's lab (University of Missouri-Columbia) to construct three normalized cDNA libraries. Dark-grown maize seedlings with primary roots 12-20 mm in length were transplanted to high (-0.03 MPa) or low water potential (-1.6 MPa) vermiculite, and harvested at 5 h and 48 h after transplanting. About 1,000 roots were used for each of the low water potential libraries (zmrw05 and zmrw48) while 500 roots were combined from each of the two time points at high water potential (zmrw00). Each root was divided into 4 segments (distances are from the junction of the root apex and root cap): segment 1, 0-3 mm plus the root cap; segment 2, 3-7 mm; segment 3, 7-12 mm; segment 4, 12-20 mm. (For details of conditions see (1) with nutrient modifications as in (2)). The three normalized cDNA libraries were constructed in the lab of Hans Bohnert (University of Illinois-UC). Total RNA was extracted by the 'hot Phenol' method (Plant Molecular Biology manual, D5: 1-13, 2nd ed., 1997). This method worked in eliminating carbohydrate material present in the root tips. The integrity of the RNA was verified by denaturing agarose gels and spectrophotometry (ratio A260/280). Poly(A)+mRNA was isolated twice from total RNA using the Oligotex Direct mRNA kit (Qiagen). Poly(A)+ mRNA was converted to double-stranded cDNA and tagged by using modified Oligo(dT) primers. One of 4 sequence tags corresponding to a different segment of the root was added to the 3'-end of the modified Oligo(dT) primers, including a NotI site and used to reverse transcribe the segment-specific mRNAs into cDNAs. Each library contains all four tags. A suffix (s1, s2, s3, or s4) has been added to each sequence identifier to designate which region of the root (Root_segment 1, 2, 3, or 4) the sequence was found in based on the identification of the tag. A suffix of s0 indicates that the sequence tag, and hence the source segment, could not be identified. The double stranded cDNAs were size-selected (>450 bp). Size selected cDNAs were adaptor with EcoRI adaptors at both ends, and then digested with NotI. The cDNA was directionally cloned into EcoRI-NotI digested pBS II SK(+) phagemid vector (Stratagene) and electroporated into E.coli DH10B. The total number of white colony forming units (cfu) in the primary libraries before amplification was as follows: zmrw05: 3.37 x 10 ⁶ ; zmrw48: 4.87 x 10 ⁶ ; zmrw00: 3 x 10 ⁶ . The background of empty clones was less than 1%. Inserts ranged from ~0.5kb to >2.5 kb, as determined by		
QY	525 GTGGAGCAAGCTGTGGGAATGCCATCAGTTCGCCAAATATTCAAGGTCCATAGTGAAGCC 584		1..814 /organism="Zea mays" /mol_type="mRNA" /db_xref="taxon:4577" /clone_lib="zmrw00" /note="Samples were collected in Robert E. Sharp's lab (University of Missouri-Columbia) to construct three normalized cDNA libraries. Dark-grown maize seedlings with primary roots 12-20 mm in length were transplanted to high (-0.03 MPa) or low water potential (-1.6 MPa) vermiculite, and harvested at 5 h and 48 h after transplanting. About 1,000 roots were used for each of the low water potential libraries (zmrw05 and zmrw48) while 500 roots were combined from each of the two time points at high water potential (zmrw00). Each root was divided into 4 segments (distances are from the junction of the root apex and root cap): segment 1, 0-3 mm plus the root cap; segment 2, 3-7 mm; segment 3, 7-12 mm; segment 4, 12-20 mm. (For details of conditions see (1) with nutrient modifications as in (2)). The three normalized cDNA libraries were constructed in the lab of Hans Bohnert (University of Illinois-UC). Total RNA was extracted by the 'hot Phenol' method (Plant Molecular Biology manual, D5: 1-13, 2nd ed., 1997). This method worked in eliminating carbohydrate material present in the root tips. The integrity of the RNA was verified by denaturing agarose gels and spectrophotometry (ratio A260/280). Poly(A)+mRNA was isolated twice from total RNA using the Oligotex Direct mRNA kit (Qiagen). Poly(A)+ mRNA was converted to double-stranded cDNA and tagged by using modified Oligo(dT) primers. One of 4 sequence tags corresponding to a different segment of the root was added to the 3'-end of the modified Oligo(dT) primers, including a NotI site and used to reverse transcribe the segment-specific mRNAs into cDNAs. Each library contains all four tags. A suffix (s1, s2, s3, or s4) has been added to each sequence identifier to designate which region of the root (Root_segment 1, 2, 3, or 4) the sequence was found in based on the identification of the tag. A suffix of s0 indicates that the sequence tag, and hence the source segment, could not be identified. The double stranded cDNAs were size-selected (>450 bp). Size selected cDNAs were adaptor with EcoRI adaptors at both ends, and then digested with NotI. The cDNA was directionally cloned into EcoRI-NotI digested pBS II SK(+) phagemid vector (Stratagene) and electroporated into E.coli DH10B. The total number of white colony forming units (cfu) in the primary libraries before amplification was as follows: zmrw05: 3.37 x 10 ⁶ ; zmrw48: 4.87 x 10 ⁶ ; zmrw00: 3 x 10 ⁶ . The background of empty clones was less than 1%. Inserts ranged from ~0.5kb to >2.5 kb, as determined by		
Db	560 GTGGAGCAAGCTGTGGGAATGCCATCAGTTCGCCAAATATTCAAGGTTCACAGGAAGCC 501		1..814 /organism="Zea mays" /mol_type="mRNA" /db_xref="taxon:4577" /clone_lib="zmrw00" /note="Samples were collected in Robert E. Sharp's lab (University of Missouri-Columbia) to construct three normalized cDNA libraries. Dark-grown maize seedlings with primary roots 12-20 mm in length were transplanted to high (-0.03 MPa) or low water potential (-1.6 MPa) vermiculite, and harvested at 5 h and 48 h after transplanting. About 1,000 roots were used for each of the low water potential libraries (zmrw05 and zmrw48) while 500 roots were combined from each of the two time points at high water potential (zmrw00). Each root was divided into 4 segments (distances are from the junction of the root apex and root cap): segment 1, 0-3 mm plus the root cap; segment 2, 3-7 mm; segment 3, 7-12 mm; segment 4, 12-20 mm. (For details of conditions see (1) with nutrient modifications as in (2)). The three normalized cDNA libraries were constructed in the lab of Hans Bohnert (University of Illinois-UC). Total RNA was extracted by the 'hot Phenol' method (Plant Molecular Biology manual, D5: 1-13, 2nd ed., 1997). This method worked in eliminating carbohydrate material present in the root tips. The integrity of the RNA was verified by denaturing agarose gels and spectrophotometry (ratio A260/280). Poly(A)+mRNA was isolated twice from total RNA using the Oligotex Direct mRNA kit (Qiagen). Poly(A)+ mRNA was converted to double-stranded cDNA and tagged by using modified Oligo(dT) primers. One of 4 sequence tags corresponding to a different segment of the root was added to the 3'-end of the modified Oligo(dT) primers, including a NotI site and used to reverse transcribe the segment-specific mRNAs into cDNAs. Each library contains all four tags. A suffix (s1, s2, s3, or s4) has been added to each sequence identifier to designate which region of the root (Root_segment 1, 2, 3, or 4) the sequence was found in based on the identification of the tag. A suffix of s0 indicates that the sequence tag, and hence the source segment, could not be identified. The double stranded cDNAs were size-selected (>450 bp). Size selected cDNAs were adaptor with EcoRI adaptors at both ends, and then digested with NotI. The cDNA was directionally cloned into EcoRI-NotI digested pBS II SK(+) phagemid vector (Stratagene) and electroporated into E.coli DH10B. The total number of white colony forming units (cfu) in the primary libraries before amplification was as follows: zmrw05: 3.37 x 10 ⁶ ; zmrw48: 4.87 x 10 ⁶ ; zmrw00: 3 x 10 ⁶ . The background of empty clones was less than 1%. Inserts ranged from ~0.5kb to >2.5 kb, as determined by		
QY	585 TTCTTTCCGGGATAATGAGAGTAGTGTCTTTGAGAGATTTGCTCTCCATCGCAGCATTAGTT 644		1..814 /organism="Zea mays" /mol_type="mRNA" /db_xref="taxon:4577" /clone_lib="zmrw00" /note="Samples were collected in Robert E. Sharp's lab (University of Missouri-Columbia) to construct three normalized cDNA libraries. Dark-grown maize seedlings with primary roots 12-20 mm in length were transplanted to high (-0.03 MPa) or low water potential (-1.6 MPa) vermiculite, and harvested at 5 h and 48 h after transplanting. About 1,000 roots were used for each of the low water potential libraries (zmrw05 and zmrw48) while 500 roots were combined from each of the two time points at high water potential (zmrw00). Each root was divided into 4 segments (distances are from the junction of the root apex and root cap): segment 1, 0-3 mm plus the root cap; segment 2, 3-7 mm; segment 3, 7-12 mm; segment 4, 12-20 mm. (For details of conditions see (1) with nutrient modifications as in (2)). The three normalized cDNA libraries were constructed in the lab of Hans Bohnert (University of Illinois-UC). Total RNA was extracted by the 'hot Phenol' method (Plant Molecular Biology manual, D5: 1-13, 2nd ed., 1997). This method worked in eliminating carbohydrate material present in the root tips. The integrity of the RNA was verified by denaturing agarose gels and spectrophotometry (ratio A260/280). Poly(A)+mRNA was isolated twice from total RNA using the Oligotex Direct mRNA kit (Qiagen). Poly(A)+ mRNA was converted to double-stranded cDNA and tagged by using modified Oligo(dT) primers. One of 4 sequence tags corresponding to a different segment of the root was added to the 3'-end of the modified Oligo(dT) primers, including a NotI site and used to reverse transcribe the segment-specific mRNAs into cDNAs. Each library contains all four tags. A suffix (s1, s2, s3, or s4) has been added to each sequence identifier to designate which region of the root (Root_segment 1, 2, 3, or 4) the sequence was found in based on the identification of the tag. A suffix of s0 indicates that the sequence tag, and hence the source segment, could not be identified. The double stranded cDNAs were size-selected (>450 bp). Size selected cDNAs were adaptor with EcoRI adaptors at both ends, and then digested with NotI. The cDNA was directionally cloned into EcoRI-NotI digested pBS II SK(+) phagemid vector (Stratagene) and electroporated into E.coli DH10B. The total number of white colony forming units (cfu) in the primary libraries before amplification was as follows: zmrw05: 3.37 x 10 ⁶ ; zmrw48: 4.87 x 10 ⁶ ; zmrw00: 3 x 10 ⁶ . The background of empty clones was less than 1%. Inserts ranged from ~0.5kb to >2.5 kb, as determined by		
Db	500 TTCTTTCCGGGATAATGAGAGTAGTGTCTTTGAGGATCTGCTCCATCGACGATTAGTT 441		1..814 /organism="Zea mays" /mol_type="mRNA" /db_xref="taxon:4577" /clone_lib="zmrw00" /note="Samples were collected in Robert E. Sharp's lab (University of Missouri-Columbia) to construct three normalized cDNA libraries. Dark-grown maize seedlings with primary roots 12-20 mm in length were transplanted to high (-0.03 MPa) or low water potential (-1.6 MPa) vermiculite, and harvested at 5 h and 48 h after transplanting. About 1,000 roots were used for each of the low water potential libraries (zmrw05 and zmrw48) while 500 roots were combined from each of the two time points at high water potential (zmrw00). Each root was divided into 4 segments (distances are from the junction of the root apex and root cap): segment 1, 0-3 mm plus the root cap; segment 2, 3-7 mm; segment 3, 7-12 mm; segment 4, 12-20 mm. (For details of conditions see (1) with nutrient modifications as in (2)). The three normalized cDNA libraries were constructed in the lab of Hans Bohnert (University of Illinois-UC). Total RNA was extracted by the 'hot Phenol' method (Plant Molecular Biology manual, D5: 1-13, 2nd ed., 1997). This method worked in eliminating carbohydrate material present in the root tips. The integrity of the RNA was verified by denaturing agarose gels and spectrophotometry (ratio A260/280). Poly(A)+mRNA was isolated twice from total RNA using the Oligotex Direct mRNA kit (Qiagen). Poly(A)+ mRNA was converted to double-stranded cDNA and tagged by using modified Oligo(dT) primers. One of 4 sequence tags corresponding to a different segment of the root was added to the 3'-end of the modified Oligo(dT) primers, including a NotI site and used to reverse transcribe the segment-specific mRNAs into cDNAs. Each library contains all four tags. A suffix (s1, s2, s3, or s4) has been added to each sequence identifier to designate which region of the root (Root_segment 1, 2, 3, or 4) the sequence was found in based on the identification of the tag. A suffix of s0 indicates that the sequence tag, and hence the source segment, could not be identified. The double stranded cDNAs were size-selected (>450 bp). Size selected cDNAs were adaptor with EcoRI adaptors at both ends, and then digested with NotI. The cDNA was directionally cloned into EcoRI-NotI digested pBS II SK(+) phagemid vector (Stratagene) and electroporated into E.coli DH10B. The total number of white colony forming units (cfu) in the primary libraries before amplification was as follows: zmrw05: 3.37 x 10 ⁶ ; zmrw48: 4.87 x 10 ⁶ ; zmrw00: 3 x 10 ⁶ . The background of empty clones was less than 1%. Inserts ranged from ~0.5kb to >2.5 kb, as determined by		
QY	645 GTTGCCACCGAGGTGGTCTGTATCCGACCAATTAATCTGGAGATATATGAAGAGGGGC 704		1..814 /organism="Zea mays" /mol_type="mRNA" /db_xref="taxon:4577" /clone_lib="zmrw00" /note="Samples were collected in Robert E. Sharp's lab (University of Missouri-Columbia) to construct three normalized cDNA libraries. Dark-grown maize seedlings with primary roots 12-20 mm in length were transplanted to high (-0.03 MPa) or low water potential (-1.6 MPa) vermiculite, and harvested at 5 h and 48 h after transplanting. About 1,000 roots were used for each of the low water potential libraries (zmrw05 and zmrw48) while 500 roots were combined from each of the two time points at high water potential (zmrw00). Each root was divided into 4 segments (distances are from the junction of the root apex and root cap): segment 1, 0-3 mm plus the root cap; segment 2, 3-7 mm; segment 3, 7-12 mm; segment 4, 12-20 mm. (For details of conditions see (1) with nutrient modifications as in (2)). The three normalized cDNA libraries were constructed in the lab of Hans Bohnert (University of Illinois-UC). Total RNA was extracted by the 'hot Phenol' method (Plant Molecular Biology manual, D5: 1-13, 2nd ed., 1997). This method worked in eliminating carbohydrate material present in the root tips. The integrity of the RNA was verified by denaturing agarose gels and spectrophotometry (ratio A260/280). Poly(A)+mRNA was isolated twice from total RNA using the Oligotex Direct mRNA kit (Qiagen). Poly(A)+ mRNA was converted to double-stranded cDNA and tagged by using modified Oligo(dT) primers. One of 4 sequence tags corresponding to a different segment of the root was added to the 3'-end of the modified Oligo(dT) primers, including a NotI site and used to reverse transcribe the segment-specific mRNAs into cDNAs. Each library contains all four tags. A suffix (s1, s2, s3, or s4) has been added to each sequence identifier to designate which region of the root (Root_segment 1, 2, 3, or 4) the sequence was found in based on the identification of the tag. A suffix of s0 indicates that the sequence tag, and hence the source segment, could not be identified. The double stranded cDNAs were size-selected (>450 bp). Size selected cDNAs were adaptor with EcoRI adaptors at both ends, and then digested with NotI. The cDNA was directionally cloned into EcoRI-NotI digested pBS II SK(+) phagemid vector (Stratagene) and electroporated into E.coli DH10B. The total number of white colony forming units (cfu) in the primary libraries before amplification was as follows: zmrw05: 3.37 x 10 ⁶ ; zmrw48: 4.87 x 10 ⁶ ; zmrw00: 3 x 10 ⁶ . The background of empty clones was less than 1%. Inserts ranged from ~0.5kb to >2.5 kb, as determined by		
Db	440 GTTGCCACCGAGGTGGTCTGTATCCGACCAAGTTAACTGGAATATATGAAGAGGGC 381		1..814 /organism="Zea mays" /mol_type="mRNA" /db_xref="taxon:4577" /clone_lib="zmrw00" /note="Samples were collected in Robert E. Sharp's lab (University of Missouri-Columbia) to construct three normalized cDNA libraries. Dark-grown maize seedlings with primary roots 12-20 mm in length were transplanted to high (-0.03 MPa) or low water potential (-1.6 MPa) vermiculite, and harvested at 5 h and 48 h after transplanting. About 1,000 roots were used for each of the low water potential libraries (zmrw05 and zmrw48) while 500 roots were combined from each of the two time points at high water potential (zmrw00). Each root was divided into 4 segments (distances are from the junction of the root apex and root cap): segment 1, 0-3 mm plus the root cap; segment 2, 3-7 mm; segment 3, 7-12 mm; segment 4, 12-20 mm. (For details of conditions see (1) with nutrient modifications as in (2)). The three normalized cDNA libraries were constructed in the lab of Hans Bohnert (University of Illinois-UC). Total RNA was extracted by the 'hot Phenol' method (Plant Molecular Biology manual, D5: 1-13, 2nd ed., 1997). This method worked in eliminating carbohydrate material present in the root tips. The integrity of the RNA was verified by denaturing agarose gels and spectrophotometry (ratio A260/280). Poly(A)+mRNA was isolated twice from total RNA using the Oligotex Direct mRNA kit (Qiagen). Poly(A)+ mRNA was converted to double-stranded cDNA and tagged by using modified Oligo(dT) primers. One of 4 sequence tags corresponding to a different segment of the root was added to the 3'-end of the modified Oligo(dT) primers, including a NotI site and used to reverse transcribe the segment-specific mRNAs into cDNAs. Each library contains all four tags. A suffix (s1, s2, s3, or s4) has been added to each sequence identifier to designate which region of the root (Root_segment 1, 2, 3, or 4) the sequence was found in based on the identification of the tag. A suffix of s0 indicates that the sequence tag, and hence the source segment, could not be identified. The double stranded cDNAs were size-selected (>450 bp). Size selected cDNAs were adaptor with EcoRI adaptors at both ends, and then digested with NotI. The cDNA was directionally cloned into EcoRI-NotI digested pBS II SK(+) phagemid vector (Stratagene) and electroporated into E.coli DH10B. The total number of white colony forming units (cfu) in the primary libraries before amplification was as follows: zmrw05: 3.37 x 10 ⁶ ; zmrw48: 4.87 x 10 ⁶ ; zmrw00: 3 x 10 ⁶ . The background of empty clones was less than 1%. Inserts ranged from ~0.5kb to >2.5 kb, as determined by		
QY	705 CTATCTGTTTGGTATAGTCCCTTGGATGCTCTTGCTAGGCGTATTGCTTAAGTGGGA 764		1..814 /organism="Zea mays" /mol_type="mRNA" /db_xref="taxon:4577" /clone_lib="zmrw00" /note="Samples were collected in Robert E. Sharp's lab (University of Missouri-Columbia) to construct three normalized cDNA libraries. Dark-grown maize seedlings with primary roots 12-20 mm in length were transplanted to high (-0.03 MPa) or low water potential (-1.6 MPa) vermiculite, and harvested at 5 h and 48 h after transplanting. About 1,000 roots were used for each of the low water potential libraries (zmrw05 and zmrw48) while 500 roots were combined from each of the two time points at high water potential (zmrw00). Each root was divided into 4 segments (distances are from the junction of the root apex and root cap): segment 1, 0-3 mm plus the root cap; segment 2, 3-7 mm; segment 3, 7-12 mm; segment 4, 12-20 mm. (For details of conditions see (1) with nutrient modifications as in (2)). The three normalized cDNA libraries were constructed in the lab of Hans Bohnert (University of Illinois-UC). Total RNA was extracted by the 'hot Phenol' method (Plant Molecular Biology manual, D5: 1-13, 2nd ed., 1997). This method worked in eliminating carbohydrate material present in the root tips. The integrity of the RNA was verified by denaturing agarose gels and spectrophotometry (ratio A260/280). Poly(A)+mRNA was isolated twice from total RNA using the Oligotex Direct mRNA kit (Qiagen). Poly(A)+ mRNA was converted to double-stranded cDNA and tagged by using modified Oligo(dT) primers. One of 4 sequence tags corresponding to a different segment of the root was added to the 3'-end of the modified Oligo(dT) primers, including a NotI site and used to reverse transcribe the segment-specific mRNAs into cDNAs. Each library contains all four tags. A suffix (s1, s2, s3, or s4) has been added to each sequence identifier to designate which region of the root (Root_segment 1, 2, 3, or 4) the sequence was found in based on the identification of the tag. A suffix of s0 indicates that the sequence tag, and hence the source segment, could not be identified. The double stranded cDNAs were size-selected (>450 bp). Size selected cDNAs were adaptor with EcoRI adaptors at both ends, and then digested with NotI. The cDNA was directionally cloned into EcoRI-NotI digested pBS II SK(+) phagemid vector (Stratagene) and electroporated into E.coli DH10B. The total number of white colony forming units (cfu) in the primary libraries before amplification was as follows: zmrw05: 3.37 x 10 ⁶ ; zmrw48: 4.87 x 10 ⁶ ; zmrw00: 3 x 10 ⁶ . The background of empty clones was less than 1%. Inserts ranged from ~0.5kb to >2.5 kb, as determined by		
Db	380 CTATCCGTTTGGTATAGTGGCCCTTGGATGCTCTTGCTAGGCGTATTGCTTAAGTGGGA 321		1..814 /organism="Zea mays" /mol_type="mRNA" /db_xref="taxon:4577" /clone_lib="zmrw00" /note="Samples were collected in Robert E. Sharp's lab (University of Missouri-Columbia) to construct three normalized cDNA libraries. Dark-grown maize seedlings with primary roots 12-20 mm in length were transplanted to high (-0.03 MPa) or low water potential (-1.6 MPa) vermiculite, and harvested at 5 h and 48 h after transplanting. About 1,000 roots were used for each of the low water potential libraries (zmrw05 and zmrw48) while 500 roots were combined from each of the two time points at high water potential (zmrw00). Each root was divided into 4 segments (distances are from the junction of the root apex and root cap): segment 1, 0-3 mm plus the root cap; segment 2, 3-7 mm; segment 3, 7-12 mm; segment 4, 12-20 mm. (For details of conditions see (1) with nutrient modifications as in (2)). The three normalized cDNA libraries were constructed in the lab of Hans Bohnert (University of Illinois-UC). Total RNA was extracted by the 'hot Phenol' method (Plant Molecular Biology manual, D5: 1-13, 2nd ed., 1997). This method worked in eliminating carbohydrate material present in the root tips. The integrity of the RNA was verified by denaturing agarose gels and spectrophotometry (ratio A260/280). Poly(A)+mRNA was isolated twice from total RNA using the Oligotex Direct mRNA kit (Qiagen). Poly(A)+ mRNA was converted to double-stranded cDNA and tagged by using modified Oligo(dT) primers. One of 4 sequence tags corresponding to a different segment of the root was added to the 3'-end of the modified Oligo(dT) primers, including a NotI site and used to reverse transcribe the segment-specific mRNAs into cDNAs. Each library contains all four tags. A suffix (s1, s2, s3, or s4) has been added to each sequence identifier to designate which region of the root (Root_segment 1, 2, 3, or 4) the sequence was found in based on the identification of the tag. A suffix of s0 indicates that the sequence tag, and hence the source segment, could not be identified. The double stranded cDNAs were size-selected (>450 bp). Size selected cDNAs were adaptor with EcoRI adaptors at both ends, and then digested with NotI. The cDNA was directionally cloned into EcoRI-NotI digested pBS II SK(+) phagemid vector (Stratagene) and electroporated into E.coli DH10B. The total number of white colony forming units (cfu) in the primary libraries before amplification was as follows: zmrw05: 3.37 x 10 ⁶ ; zmrw48: 4.87 x 10 ⁶ ; zmrw00: 3 x 10 ⁶ . The background of empty clones was less than 1%. Inserts ranged from ~0.5kb to >2.5 kb, as determined by		
QY	765 ACTGCTCTCGTCTCTTCTGGACCAACCATCTGCTGATCCGTAGCGCAATGGCCCTTTCT 824		1..814 /organism="Zea mays" /mol_type="mRNA" /db_xref="taxon:4577" /clone_lib="zmrw00" /note="Samples were collected in Robert E. Sharp's lab (University of Missouri-Columbia) to construct three normalized cDNA libraries. Dark-grown maize seedlings with primary roots 12-20 mm in length were transplanted to high (-0.03 MPa) or low water potential (-1.6 MPa) vermiculite, and harvested at 5 h and 48 h after transplanting. About 1,000 roots were used for each of the low water potential libraries (zmrw05 and zmrw48) while 500 roots were combined from each of the two time points at high water potential (zmrw00). Each root was divided into 4 segments (distances are from the junction of the root apex and root cap): segment 1, 0-3 mm plus the root cap; segment 2, 3-7 mm; segment 3, 7-12 mm; segment 4, 12-20 mm. (For details of conditions see (1) with nutrient modifications as in (2)). The three normalized cDNA libraries were constructed in the lab of Hans Bohnert (University of Illinois-UC). Total RNA was extracted by the 'hot Phenol' method (Plant Molecular Biology manual, D5: 1-13, 2nd ed., 1997). This method worked in eliminating carbohydrate material present in the root tips. The integrity of the RNA was verified by denaturing agarose gels and spectrophotometry (ratio A260/280). Poly(A)+mRNA was isolated twice from total RNA using the Oligotex Direct mRNA kit (Qiagen). Poly(A)+ mRNA was converted to double-stranded cDNA and tagged by using modified Oligo(dT) primers. One of 4 sequence tags corresponding to a different segment of the root was added to the 3'-end of the modified Oligo(dT) primers, including a NotI site and used to reverse transcribe the segment-specific mRNAs into cDNAs. Each library contains all four tags. A suffix (s1, s2, s3, or s4) has been added to each sequence identifier to designate which region of the root (Root_segment 1, 2, 3, or 4) the sequence was found in based on the identification of the tag. A suffix of s0 indicates that the sequence tag, and hence the source segment, could not be identified. The double stranded cDNAs were size-selected (>450 bp). Size selected cDNAs were adaptor with EcoRI adaptors at both ends, and then digested with NotI. The cDNA was directionally cloned into EcoRI-NotI digested pBS II SK(+) phagemid vector (Stratagene) and electroporated into E.coli DH10B. The total number of white colony forming units (cfu) in the primary libraries before amplification was as follows: zmrw05: 3.37 x 10 ⁶ ; zmrw48: 4.87 x 10 ⁶ ; zmrw00: 3 x 10 ⁶ . The background of empty clones was less than 1%. Inserts ranged from ~0.5kb to >2.5 kb, as determined by		
Db	320 ACCGCTTCTCGTCTCTTCTGGACCAACCGTGGTATCATACAAATGGCCTTTCT 261		1..814 /organism="Zea mays" /mol_type="mRNA" /db_xref="taxon:4577" /clone_lib="zmrw00" /note="Samples were collected in Robert E. Sharp's lab (University of Missouri-Columbia) to construct three normalized cDNA libraries. Dark-grown maize seedlings with primary roots 12-20 mm in length were transplanted to high (-0.03 MPa) or low water potential (-1.6 MPa) vermiculite, and harvested at 5 h and 48 h after transplanting. About 1,000 roots were used for each of the low water potential libraries (zmrw05 and zmrw48) while 500 roots were combined from each of the two time points at high water potential (zmrw00). Each root was divided into 4 segments (distances are from the junction of the root apex and root cap): segment 1, 0-3 mm plus the root cap; segment 2, 3-7 mm; segment 3, 7-12 mm; segment 4, 12-20 mm. (For details of conditions see (1) with nutrient modifications as in (2)). The three normalized cDNA libraries were constructed in the lab of Hans Bohnert (University of Illinois-UC). Total RNA was extracted by the 'hot Phenol' method (Plant Molecular Biology manual, D5: 1-13, 2nd ed., 1997). This method worked in eliminating carbohydrate material present in the root tips. The integrity of the RNA was verified by denaturing agarose gels and spectrophotometry (ratio A260/280). Poly(A)+mRNA was isolated twice from total RNA using the Oligotex Direct mRNA kit (Qiagen). Poly(A)+ mRNA was converted to double-stranded cDNA and tagged by using modified Oligo(dT) primers. One of 4 sequence tags corresponding to a different segment of the root was added to the 3'-end of the modified Oligo(dT) primers, including a NotI site and used to reverse transcribe the segment-specific mRNAs into cDNAs. Each library contains all four tags. A suffix (s1, s2, s3, or s4) has been added to each sequence identifier to designate which region of the root (Root_segment 1, 2, 3, or 4) the sequence was found in based on the identification of the tag. A suffix of s0 indicates that the sequence tag, and hence the source segment, could not be identified. The double stranded cDNAs were size-selected (>450 bp). Size selected cDNAs were adaptor with EcoRI adaptors at both ends, and then digested with NotI. The cDNA was directionally cloned into EcoRI-NotI digested pBS II SK(+) phagemid vector (Stratagene) and electroporated into E.coli DH10B. The total number of white colony forming units (cfu) in the primary libraries before amplification was as follows: zmrw05: 3.37 x 10 ⁶ ; zmrw48: 4.87 x 10 ⁶ ; zmrw00: 3 x 10 ⁶ . The background of empty clones was less than 1%. Inserts ranged from ~0.5kb to >2.5 kb, as determined by		
QY	825 AAGCTCAGCATGTTGCACAGCAAGGGGTGATGCTTATGCAATGCGAGATGTAAGGGTT 884		1..814 /organism="Zea mays" /mol_type="mRNA" /db_xref="taxon:4577" /clone_lib="zmrw00" /note="Samples were collected in Robert E. Sharp's lab (University of Missouri-Columbia) to construct three normalized cDNA libraries. Dark-grown maize seedlings with primary roots 12-20 mm in length were transplanted to high (-0.03 MPa) or low water potential (-1.6 MPa) vermiculite, and harvested at 5 h and 48 h after transplanting. About 1,000 roots were used for each of the low water potential libraries (zmrw05 and zmrw48) while 500 roots were combined from each of the two time points at high water potential (zmrw00). Each root was divided into 4 segments (distances are from the junction of the root apex and root cap): segment 1, 0-3 mm plus the root cap; segment 2, 3-7 mm; segment 3, 7-12 mm; segment 4, 12-20 mm. (For details of conditions see (1) with nutrient modifications as in (2)). The three normalized cDNA libraries were constructed in the lab of Hans Bohnert (University of Illinois-UC). Total RNA was extracted by the 'hot Phenol' method (Plant Molecular Biology manual, D5: 1-13, 2nd ed., 1997). This method worked in eliminating carbohydrate material present in the root tips. The integrity of the RNA was verified by denaturing agarose gels and spectrophotometry (ratio A260/280). Poly(A)+mRNA was isolated twice from total RNA using the Oligotex Direct mRNA kit (Qiagen). Poly(A)+ mRNA was converted to double-stranded cDNA and tagged by using modified Oligo(dT) primers. One of 4 sequence tags corresponding to a different segment of the root was added to the 3'-end of the modified Oligo(dT) primers, including a NotI site and used to reverse transcribe the segment-specific mRNAs into cDNAs. Each library contains all four tags. A suffix (s1, s2, s3, or s4) has been added to each sequence identifier to designate which region of the root (Root_segment 1, 2, 3, or 4) the sequence was found in based on the identification of the tag		

PCR. Plasmid DNA from the primary libraries then was converted to single-stranded circles and used as a template for PCR amplification using the T7 and T3 priming sites that flank the cloned cDNA inserts. The purified PCR products, representing the entire cDNA population cloned in each library, were used as a driver for normalization. Hybridization between the single-stranded library and the PCR products was carried out for 44 hours at 30°C. Non-hybridized single-stranded DNA circles were separated from hybridized DNA rendered partially double-stranded and electrophoreted into DH10B. The total number of clones with insert was: zmrws05: 2.0x10⁷; zmrws48: 4.2x10⁷; zmrws00: 1.1x10⁷. The background of empty clones was less than 2%. Insert size, determined by PCR of the entire library, ranged from 0.5kb to 2.5kb. (1) Sharp R E; Silk W K; Heiao T C. Growth of the Maize Primary Root at Low Water Potentials I. Spatial Distribution of Expansive Growth. Plant Physiology (Rockville). 87(1). 1988. 50-57. (2) Spollen W G; LeNoble M E; Samuels T D; Bernstein N; Sharp R E. Abscissic acid accumulation maintains maize primary root elongation at low water potentials by restricting ethylene production. Plant Physiology (Rockville). 122(3). March, 2000. 967-976.
TAG TISSUE=Root _segment_3
TAG_SEQ=TCGCA"

ORIGIN

Query Match 44.8%; Score 537.2; DB 7; Length 814;
Best Local Similarity 92.2%; Pred. No. 4.6e-119;
Matches 566; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

QY 472 GGAAGATTATGCTGAAGTCTGGGTATTCGTTCTTGATAGTGACAAATTAGTGGAGC 531
DB 806 GGGAGATCATGCTGAAGTCTGGGTATTCGTTCTTGATAGTGACAAATTAGTGGAGC 747

QY 532 AAGCTGTTGGAATGCCATCAGTTGCCAAATATTCAGGTCATAGTGAAGCTCTCTTTC 591
DB 746 AAGCTGTTGGAATGCCATCAGTTGCCAAATATTCAGGTCATAGTGAAGCTCTCTTTC 687

QY 592 GGGATTAATGAGAGTAGTGTCTTGAGAGATTGTCCTCCATGCGACGATTAGTTGGCA 651
DB 686 GGGATTAATGAGAGTAGTGTCTTGAGAGATTGTCCTCCATGCGACGATTAGTTGGCA 627

QY 652 CCGGAGTGGTCTGTTATCCGACCAATTAACGTGGAGATATAGAGAGGCGCTATCTG 711
DB 626 CCGGAGTGGTCTGTTATCCGACCAATTAACGTGGAGATATAGAGAGGCGCTATCTG 567

QY 712 TTTGGTTAGATGTCCTTGGATGCTCTTGCTAGGCGTATTGCTAAAGTGGCAATGCTCT 771
DB 566 TTTGGTTAGATGTCCTTGGATGCTCTTGCTAGGCGTATTGCTAAAGTGGCAATGCTCT 507

QY 772 CTGCTCCTCTTTGGACCAACCAATCTGGTGATCCGTPACGCAATGGCCTTTTCTAAGCTCA 831
DB 506 CTGCTCCTCTTTGGACCAACCAATCTGGTGATCCGTPACGCAATGGCCTTTTCTAAGCTCA 447

QY 832 GCATGCTCTCCAGCAAGGGGTGATGCTATGCAAAATGCAATGCAATGCAATGCTCTGCG 891
DB 446 GCATGCTCTCCAGCAAGGGGTGATGCTATGCAAAATGCAATGCAATGCAATGCTCTGCG 387

QY 892 AAGAGATTGCATGTAACCAAGGTTCATGATGATGCTCTTAAGCTGACACCTACTGATATTG 951
DB 386 AAGAGATTGCATGTAACCAAGGTTCATGATGATGCTCTTAAGCTGATGATGATGATGATG 327

QY 952 CAATTGAGTCACTTCATAGATGAGAGTTCGTCTATCGAGCACACTGCTGATGATTCAG 1011
DB 326 CAATTGAGTCACTTCATAGATGAGAGTTCGTCTATCGAGCACACTGCTGATGATTCAG 267

QY 1012 CTAGCCACGGCAAGCTGAGTGCAGATCCAGAGGATACAGACCTTGTAAGACCTTAATC 1071
DB 266 CTAGCCACCTGCAAGCTGAGTGCAGATCCAGAGGATACAGACCTTGTAAGACCTTAATC 207

QY 1072 CCTTTGTTGGCCAC 1085
DB 1072 CCTTTGTTGGCCAC 1085

DB 206 CTTCTGTTTTGTAC 193

RESULT 13
CA105686
LOCUS
DEFINITION
SCJFHRIC08D05.g HR1 Saccharum officinarum cDNA clone SCJFHRIC08D05
5', mRNA sequence.
CA105686 650 bp mRNA linear EST 23-SEP-2003
SCJFHRIC08D05
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Saccharum officinarum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Saccharum; Saccharum officinarum
complex.
1 (bases 1 to 650)
Vettore,A.L., da Silva,P.R., Kemper,E.L. and Arruda,P.
The libraries that made SUCST
Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
Contact: Arruda P
Centro de Biologia Molecular e Engenharia Genetica
Universidade Estadual de Campinas
Caixa Postal 6010, 13083-970, Campinas SP, Brazil
Tel: 55 19 3788 1137
Fax: 55 19 3788 1089
Email: parruda@unicamp.br
Clone distribution: clone distribution information can be found
through the Brazilian Clone Collection Center (BCCC) at
http://www.bccccenter.fcav.unesp.br
Plate: C08 Row: D Column: 05
Seq primer: 17 Promoter Primer.
Location/Qualifiers
1..650
/organism="Saccharum officinarum"
/mol_type="mRNA"
/db_xref="taxon:4547"
/clone="SCJFHRIC08D05"
/lab_host="DH10B"
/clone_lib="HR1"
/note="Organ: seedlings inoculated with Herbaspirillum
rubrisubalbicans; Vector: pSport1; Site 1: SalI; Site 2:
NotI; An unidirectional cDNA library generated from
rubrisubalbicans; cDNA was prepared from polyA+ mRNA
using SuperScript Plasmid System Kit (Invitrogen). The
double-strand cDNAs were fractionated in a sepharose
CL-2B 40cm-columns and fragments sizing between 0.8 and
1.5 Kb were directionally cloned into the vector. Details
of each source of RNA and library construction can be
obtained at http://sucst.lad.ic.unicamp.br/public"

ORIGIN
Query Match 44.2%; Score 530.8; DB 6; Length 650;
Best Local Similarity 93.0%; Pred. No. 1.6e-117;
Matches 569; Conservative 0; Mismatches 37; Indels 6; Gaps 1;

QY 130 CCGCGAGCGCAACAAATGGAGCGGGGGCGTGGCTCGCGCTGAGCGGGGGCGG 189
DB 45 CCGCGAGCGCAACAAATGGAGCGGGGGCGTGGCTCGCGCTGAGCGGGGGCGG 104

QY 190 GCTTCGGTCCAGCCGACCGGGGGCGCTTACAGCGGCCACCGGAGCCTGAGAGTCG 249
DB 105 GCTTCGGTCCGCGCGCGCGGGCGCTTACAGTCGCCCGCGGAGCCTCAGAGTCG 164

QY 250 CTGACCCGCGGACCTCGGTCGCTGTGCGGGCTCGGGGTCCAAAGCCCGTCGACCCG 309
DB 165 CTGCGCCGCTCGGAGCAGCGGTTGCTGTGCGGGCTCGGGGTCCAAAGCCCGTCGACCG- 223

QY 310 TCCGACTCCGTGCGAAGAAATCGTCCGAGGTCATGAAACATCGCACAACTCGTTGAGC 369
DB 224 -----CTCCGTGCGAAGAAATCGTTCGGTGGTTCATGAAACATTCGATAACTCCGTTGAGC 278

LOCUS CB684026 818 bp mRNA linear EST 09-APR-2003
DEFINITION OSJNEF13C12_r OSJNEF Oryza sativa (japonica cultivar-group) cDNA
clone OSJNEF13C12 3', mRNA sequence.
ACCESSION CB684026
VERSION
KEYWORDS
SOURCE
ORGANISM
Oryza sativa (japonica cultivar-group)
Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.

REFERENCE 1 (bases 1 to 818)
Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E.,
Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.
TITLE Large-scale identification of ESTs involved in the interaction
between rice and Magnaporthe grisea
JOURNAL Unpublished (2003)

COMMENT Contact: Rod Wing
Arizona Genomics Institute
University of Arizona
Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
85721-0088, USA
Tel: 520 626 3967
Fax: 520 621 9288
Email: http://genome.arizona.edu

PCR Primers
FORWARD: gta aaa cga cgg cca gtcg
BACKWARD: gga aac agc tat gac cat g
Plate: 13 row: C column: 12
Seq primer: gga aac agc tat gac cat g.

FEATURES
Location/Qualifiers
1..818
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="OSJNEF13C12"
/tissue_type="Leaf"
/dev_stage="3 week"
/lab_host="DH10B"
/clone_lib="OSJNEF"
/note="Vector: pBluescript II KS +; Site_1: EcoRI; Site_2:
XhoI; Uninfected Control"

ORIGIN
Query Match 43.4%; Score 520.8; DB 6; Length 818;
Best Local Similarity 79.7%; Pred. No. 4.3e-115;
Matches 630; Conservative 0; Mismatches 152; Indels 8; Gaps 1;

QY 417 AACGGAGGTATTTTACCTAGTAGGAATGATGGGTTCTGGAAAAAGTACTGTGGGGAAG 476
DB 818 AATGGACGGTGTATTTTACCTAGTTGGAAATGATGGGTTCTGGAAAAAGTACTGTGGGGAAG 759

QY 477 ATTATCTCTGAAGTCTTGGGTTATTCGTTCTTTGATAGTGACAAAGTTAGTGGAGCAAGCT 536
DB 758 ATCATGCTGAAGCTTTGGGTTATTCGTTCTTTGATAGTAAATTTGGTCGAACAAAGCT 699

QY 537 GTTGGAAATGCCATCAGTTGCCAAATATTCAGGTCATAGTGAAGCCTTCTTTGGGAT 596
DB 698 GTGGGATGCTCTTCAAGTCGCTCAAAATTTTCAAGGTTTCATAGTGAAGCCTTCTTTAGGGAT 639

QY 597 AATGACAGTAGTCTTTGAGAGATTGTCTCCATCGACGATTTAGTTGTCACCGGA 656
DB 638 AATGACAGTAGTCTTTGAGGATTGTCTCTCATGAGGATTAGTTGTTGCTACTGGA 579

QY 657 GGTGGTCTGTATTCGACCAATTAACCTGGAGATATATGAAGAGGGGCCCTATCTGTTGG 716
DB 578 GGTGGTCTGTATTCGACCAATTAACCTGGAAATACATGAAGAAGGGCCCTATCTGTTGG 519

QY 717 TTAGATGTGCCCTTGGATGCTCTTGCTAGCGTATTTGCTAAAGTGGGAACCTGCTCTCGT 776
DB 518 TTGGATGTGCCCTTGGACGCTCTTGCTAGCGTATTTGCTAAAGTGGGGAACCTGCTCTCCGT 459

QY 777 CCTCTTCTGGACCAACCACTCTGGTGTATCCGTACGCAATGGCGCTTTCTAAGCTCAGCATG 836
DB 458 CCTCTTCTTAGATCAACCACTCTGGTGTATCCATACACAATGGCTTTTCTAAGCTCAGCATG 399
QY 837 CTTGCACAGCAAAAGGGGTGATGCTTTATGCAAAATGCAGATGTAAGGGTTTCTCTGGAAGAG 896
DB 398 CTCGCGGAGCAAAAGGGCGGATGCTTTATGCAAAATGCTGATGTGAGGGTTTCTCTTGAAGAG 339
QY 897 ATTGCATGTAAACAAGGTTCATGATGATGCTCTTAGCTGACACCTACTGATATTCGAATT 956
DB 338 ATTGCATGTAAACAAGGTTCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 279
QY 957 GAGTCACCTTCAATAGATCGAGAGCTTCGTCATCGAGCACACTGCTGATAGTTTCAGCTAGC 1016
DB 278 GAGTCGTTTCATAGATCGAGAGCTTCGTCATCGAGCACACTGCTGATAGTTTCAGCTAGC 219
QY 1017 GACGCGCAAGCTGAGTCCGAGATCCAGAGGATACAGACCTTGTAGAACCTTAAATCCCTTT 1076
DB 218 GACTCCAGGCTGACTCACGCTCTCAGAGGATACAGAGCTTGTAGAAATTTTGTATCTTTT 159
QY 1077 GTTTGC-----CACATAGAGCATCGTTGAGTTATTTGTAAAGGAATGGAAGAGGA 1128
DB 158 TGTACCTTAGTGTACCTTGTAGCGCTGTGAGTTATTTGTTGCTGTACCGACAGATG 99
QY 1129 GCTAATAATCCGAAGTGTGCGCTTGGCTCAAAAAAAGGAGTATTTTCTTGTAAACTGTAAAGAAAGGATATATGAGGT 39
DB 98 ATGAAAAGAACCGGAAGTGTATTTTCTTGTAAACTGTAAAGAAAGGATATATGAGGT 39
QY 1189 AAAAAA 1198
DB 38 AATCAGAAAA 29

Search completed: August 25, 2005, 20:22:24
Job time : 4431 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 25, 2005, 10:51:49 ; Search time 167 Seconds
(without alignments)
706.359 Million cell updates/sec

Title: US-10-660-226-10
Perfect score: 1526
Sequence: 1 MEAGGVGLALQARAAGFGSS.....ADSSASDAQESQIQRIQL 305

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A Geneseq_16Dec04:.*
1: Geneseqp1980s:.*
2: Geneseqp1980s:.*
3: Geneseqp2000s:.*
4: Geneseqp2001s:.*
5: Geneseqp2002s:.*
6: Geneseqp2003as:.*
7: Geneseqp2003bs:.*
8: Geneseqp2004s:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	660	43.3	303	3 AAG49993	Aag49993 Arabidops
2	656	43.0	303	3 AAG12963	Aag12963 Arabidops
3	650	42.6	220	3 AAG50002	Aag50002 Arabidops
4	650	42.6	220	3 AAG50011	Aag50011 Arabidops
5	650	42.6	220	3 AAG50001	Aag50001 Arabidops
6	650	42.6	250	3 AAG50010	Aag50010 Arabidops
7	649.5	42.6	292	5 ABB91833	Abb91833 Herbicida
8	603	39.5	199	3 AAG49994	Aag49994 Arabidops
9	603	39.5	199	3 AAG50003	Aag50003 Arabidops
10	603	39.5	199	3 AAG50012	Aag50012 Arabidops
11	602	39.4	301	3 AAG42178	Aag42178 Arabidops
12	602	39.4	316	3 AAG42177	Aag42177 Arabidops
13	599	39.3	199	3 AAG12964	Aag12964 Arabidops
14	581	38.1	194	3 AAG49995	Aag49995 Arabidops
15	577	37.8	194	3 AAG12965	Aag12965 Arabidops
16	545	35.7	200	3 AAG42179	Aag42179 Arabidops
17	530.5	34.8	231	3 AAG27626	Aag27626 Arabidops
18	523	34.3	195	3 AAG27627	Aag27627 Arabidops
19	518	33.9	193	3 AAG12964	Aag12964 Arabidops
20	461	30.2	183	3 AAG14899	Aag14899 Arabidops
21	451	29.6	163	3 AAG23896	Aag23896 Arabidops
22	429	28.1	305	5 ABB93307	Abb93307 Herbicida
23	401	26.3	164	3 AAG14900	Aag14900 Arabidops
24	374	24.5	103	5 ABB35328	Abb35328 Human kin
25	350.5	23.0	133	3 AAG14901	Aag14901 Arabidops

26	318.5	20.9	189	8	ADM19967	Adm19967 Bacterial
27	316.5	20.7	188	8	ADG30920	Adg30920 Bacterial
28	311.5	20.4	274	5	ABB92451	Abb92451 Herbicida
29	307	20.1	176	8	ADS42564	Ads42564 Bacterial
30	297	19.5	111	3	AAG23897	Aag23897 Arabidops
31	294.5	19.3	134	8	ADS30077	Ads30077 Bacterial
32	252	16.5	194	8	ADS42142	Ads42142 Bacterial
33	249.5	16.3	168	8	ADM17512	Adm17512 Bacterial
34	243.5	16.0	231	8	ADL04401	Adl04401 M. catarr
35	237.5	15.6	492	8	ADM20281	Adm20281 Bacterial
36	237	15.5	225	7	ABM73793	Abm73793 DNA clone
37	234.5	15.4	200	8	ADS27980	Ads27980 Bacterial
38	231	15.1	180	8	ADR31461	Adr31461 Shikimate
39	228.5	15.0	168	8	ADN25346	Adn25346 Bacterial
40	227.5	14.9	165	8	ADN26123	Adn26123 Bacterial
41	222.5	14.6	166	8	ADN27066	Adn27066 Bacterial
42	222	14.5	200	6	ADA34862	Ada34862 Acinetoba
43	222	14.5	200	7	ABO64295	AbO64295 Klebsiell
44	220	14.4	146	8	ADS43278	Ads43278 Bacterial
45	220	14.4	148	8	ADS28767	Ads28767 Bacterial

ALIGNMENTS

RESULT 1
AAG49993
ID AAG49993 standard; protein; 303 AA.
XX AC AAG49993;
XX 18-OCT-2000 (first entry)
XX Arabidopsis thaliana protein fragment SEQ ID NO: 63306.
XX Arabidopsis thaliana
XX Protein identification; signal transduction pathway; metabolic pathway;
XX hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
XX Arabidopsis thaliana.
XX EP1033405-A2.
XX 06-SEP-2000.
XX 25-FEB-2000; 2008EP-00301439.
XX 25-FEB-1999; 99US-0121825P.
XX 05-MAR-1999; 99US-0123180P.
XX 09-MAR-1999; 99US-0123548P.
XX 23-MAR-1999; 99US-0125788P.
XX 25-MAR-1999; 99US-0126264P.
XX 29-MAR-1999; 99US-0126785P.
XX 01-APR-1999; 99US-0127462P.
XX 06-APR-1999; 99US-0128234P.
XX 08-APR-1999; 99US-0128714P.
XX 16-APR-1999; 99US-0129845P.
XX 19-APR-1999; 99US-0130077P.
XX 21-APR-1999; 99US-0130449P.
XX 23-APR-1999; 99US-0130510P.
XX 28-APR-1999; 99US-0130891P.
XX 30-APR-1999; 99US-0131449P.
XX 30-APR-1999; 99US-0132048P.
XX 30-APR-1999; 99US-0132407P.
XX 04-MAY-1999; 99US-0132484P.
XX 05-MAY-1999; 99US-0132485P.
XX 06-MAY-1999; 99US-0132486P.
XX 06-MAY-1999; 99US-0132487P.
XX 07-MAY-1999; 99US-0132863P.
XX 11-MAY-1999; 99US-0134256P.
XX 14-MAY-1999; 99US-0134218P.
XX 14-MAY-1999; 99US-0134219P.
XX 14-MAY-1999; 99US-0134221P.

PR	25-OCT-1999;	99US-0161406P.		PR	07-MAY-1999;	99US-0132863P.
PR	26-OCT-1999;	99US-0161159P.		PR	11-MAY-1999;	99US-0134256P.
PR	26-OCT-1999;	99US-01611360P.		PR	14-MAY-1999;	99US-0134218P.
PR	26-OCT-1999;	99US-01611361P.		PR	14-MAY-1999;	99US-0134219P.
PR	28-OCT-1999;	99US-01611920P.		PR	14-MAY-1999;	99US-0134221P.
PR	28-OCT-1999;	99US-01611922P.		PR	14-MAY-1999;	99US-0134370P.
PR	28-OCT-1999;	99US-01611933P.		PR	18-MAY-1999;	99US-0134768P.
PR	29-OCT-1999;	99US-0162142P.		PR	19-MAY-1999;	99US-0134941P.
Query Match 43.3%; Score 660; DB 3; Length 303;				PR	20-MAY-1999;	99US-0135124P.
Best Local Similarity 63.1%; Pred. No. 2.7e-60;				PR	21-MAY-1999;	99US-0135353P.
Matches 128; Conservative 35; Mismatches 40; Indels 0; Gaps 0;				PR	24-MAY-1999;	99US-0135629P.
				PR	25-MAY-1999;	99US-0136021P.
OY	79	LLKRKSEVLYFYLNGRCIYLVGMGSGKSTVGKIMSEVLGYSFFDSKLVQAVGMPSVA 138		PR	27-MAY-1999;	99US-0136392P.
DB	88	ILKRKAEEVKPYLNGRSMYLVGMGSGKTTVGKLSKLVGYTFDCDTLIEQAMNGTSA 147		PR	28-MAY-1999;	99US-0136782P.
OY	139	OIFKVHSEAFPRDNESVLRLDLSMRLLVATCGGAVIRPINWYMKRGLSVWLDVPLDA 198		PR	01-JUN-1999;	99US-0137222P.
DB	148	EIFVHGENFPRGKETDALKLSSRYQVVVSTGGGAVIRPINWYMKRGLSVWLDVPLDA 207		PR	03-JUN-1999;	99US-0137528P.
OY	199	LARRIAKVTASPLLDQPSGDPYAMAFSKLSMLAQRGDAYANADVRSLEIEACKQGH 258		PR	04-JUN-1999;	99US-0137502P.
DB	208	LAURIAAVGTDSPRLHDESGDAYSAFKRLSAIWDERGEAYTNARVSLLENIAKRGY 267		PR	07-JUN-1999;	99US-0137722P.
OY	259	DDVSKLTPTDIAIESLHKIESFV 281		PR	08-JUN-1999;	99US-0138094P.
DB	268	KNVSDLTPTTEIAIEAEQVLSFL 290		PR	10-JUN-1999;	99US-0138540P.
RESULT 2				PR	10-JUN-1999;	99US-0138847P.
AAG12963				PR	14-JUN-1999;	99US-0139119P.
XX	ID AAG12963 standard; protein; 303 AA.			PR	16-JUN-1999;	99US-0139452P.
AC	AAG12963;			PR	16-JUN-1999;	99US-0139453P.
XX				PR	17-JUN-1999;	99US-0139492P.
DT	17-OCT-2000 (first entry)			PR	18-JUN-1999;	99US-0139454P.
DE	Arabidopsis thaliana protein fragment SEQ ID NO: 12279.			PR	18-JUN-1999;	99US-0139455P.
XX				PR	18-JUN-1999;	99US-0139461P.
KW	Protein identification; signal transduction pathway; metabolic pathway;			PR	18-JUN-1999;	99US-0139462P.
KW	hybridisation assay; genetic mapping; gene expression control; promoter;			PR	18-JUN-1999;	99US-0139463P.
KW	termination sequence.			PR	18-JUN-1999;	99US-0139457P.
XX				PR	18-JUN-1999;	99US-0139458P.
OS	Arabidopsis thaliana.			PR	18-JUN-1999;	99US-0139459P.
PN	EP1033405-A2.			PR	18-JUN-1999;	99US-0139460P.
XX				PR	18-JUN-1999;	99US-0139461P.
XX				PR	18-JUN-1999;	99US-0139462P.
XX				PR	18-JUN-1999;	99US-0139463P.
XX				PR	18-JUN-1999;	99US-0139750P.
XX				PR	18-JUN-1999;	99US-0139763P.
XX				PR	21-JUN-1999;	99US-0139817P.
XX				PR	22-JUN-1999;	99US-0139899P.
XX				PR	23-JUN-1999;	99US-0140353P.
XX				PR	23-JUN-1999;	99US-0140354P.
XX				PR	24-JUN-1999;	99US-0140695P.
XX				PR	28-JUN-1999;	99US-0140823P.
XX				PR	29-JUN-1999;	99US-0140991P.
XX				PR	30-JUN-1999;	99US-0141287P.
XX				PR	01-JUL-1999;	99US-0141842P.
XX				PR	01-JUL-1999;	99US-0142154P.
XX				PR	02-JUL-1999;	99US-0142055P.
XX				PR	06-JUL-1999;	99US-0142390P.
XX				PR	08-JUL-1999;	99US-0142803P.
XX				PR	09-JUL-1999;	99US-0142920P.
XX				PR	12-JUL-1999;	99US-0142977P.
XX				PR	13-JUL-1999;	99US-0143542P.
XX				PR	14-JUL-1999;	99US-0143624P.
XX				PR	15-JUL-1999;	99US-0144005P.
XX				PR	16-JUL-1999;	99US-0144085P.
XX				PR	16-JUL-1999;	99US-0144086P.
XX				PR	19-JUL-1999;	99US-0144325P.
XX				PR	19-JUL-1999;	99US-0144331P.
XX				PR	19-JUL-1999;	99US-0144332P.
XX				PR	19-JUL-1999;	99US-0144333P.
XX				PR	19-JUL-1999;	99US-0144334P.
XX				PR	19-JUL-1999;	99US-0144335P.
XX				PR	20-JUL-1999;	99US-0144352P.
XX				PR	20-JUL-1999;	99US-0144632P.
XX				PR	20-JUL-1999;	99US-0144884P.
XX				PR	21-JUL-1999;	99US-0144814P.
XX				PR	21-JUL-1999;	99US-0145086P.
XX				PR	21-JUL-1999;	99US-0145088P.
XX				PR	22-JUL-1999;	99US-0145085P.
XX				PR	22-JUL-1999;	99US-0145087P.
XX				PR	22-JUL-1999;	99US-0145089P.
XX				PR	22-JUL-1999;	99US-0145192P.

PR 23-JUL-1999;	99US-0145145P.	PR 22-OCT-1999;	99US-0160980P.
PR 23-JUL-1999;	99US-0145218P.	PR 22-OCT-1999;	99US-0160981P.
PR 23-JUL-1999;	99US-0145224P.	PR 22-OCT-1999;	99US-0160989P.
PR 26-JUL-1999;	99US-0145276P.	PR 25-OCT-1999;	99US-0161404P.
PR 27-JUL-1999;	99US-0145913P.	PR 25-OCT-1999;	99US-0161405P.
PR 27-JUL-1999;	99US-0145918P.	PR 25-OCT-1999;	99US-0161406P.
PR 27-JUL-1999;	99US-0145919P.	PR 26-OCT-1999;	99US-0161359P.
PR 28-JUL-1999;	99US-0145951P.	PR 26-OCT-1999;	99US-0161360P.
PR 02-AUG-1999;	99US-0146386P.	PR 26-OCT-1999;	99US-0161361P.
PR 02-AUG-1999;	99US-0146388P.	PR 28-OCT-1999;	99US-0161920P.
PR 02-AUG-1999;	99US-0146389P.	PR 28-OCT-1999;	99US-0161922P.
PR 03-AUG-1999;	99US-0147038P.	PR 28-OCT-1999;	99US-0161993P.
PR 04-AUG-1999;	99US-0147204P.	PR 29-OCT-1999;	99US-0162142P.
PR 04-AUG-1999;	99US-0147302P.		
PR 05-AUG-1999;	99US-0147192P.		
PR 05-AUG-1999;	99US-0147260P.		
PR 06-AUG-1999;	99US-0147303P.		
PR 06-AUG-1999;	99US-0147416P.		
PR 09-AUG-1999;	99US-0147493P.		
PR 09-AUG-1999;	99US-0147935P.		
PR 10-AUG-1999;	99US-0148171P.		
PR 11-AUG-1999;	99US-0148319P.		
PR 12-AUG-1999;	99US-0148341P.		
PR 13-AUG-1999;	99US-0148565P.		
PR 13-AUG-1999;	99US-0148684P.		
PR 16-AUG-1999;	99US-0149368P.		
PR 17-AUG-1999;	99US-0149175P.		
PR 18-AUG-1999;	99US-0149426P.		
PR 20-AUG-1999;	99US-0149722P.		
PR 20-AUG-1999;	99US-0149723P.		
PR 20-AUG-1999;	99US-0149929P.		
PR 23-AUG-1999;	99US-0149302P.		
PR 23-AUG-1999;	99US-0149320P.		
PR 25-AUG-1999;	99US-0150566P.		
PR 26-AUG-1999;	99US-0150884P.		
PR 27-AUG-1999;	99US-0151065P.		
PR 27-AUG-1999;	99US-0151066P.		
PR 27-AUG-1999;	99US-0151080P.		
PR 30-AUG-1999;	99US-0151303P.		
PR 31-AUG-1999;	99US-0151438P.		
PR 01-SEP-1999;	99US-0151930P.		
PR 07-SEP-1999;	99US-0152363P.		
PR 10-SEP-1999;	99US-0153070P.		
PR 13-SEP-1999;	99US-0153758P.		
PR 15-SEP-1999;	99US-0154018P.		
PR 16-SEP-1999;	99US-0154039P.		
PR 20-SEP-1999;	99US-0154779P.		
PR 22-SEP-1999;	99US-0155139P.		
PR 23-SEP-1999;	99US-0155486P.		
PR 24-SEP-1999;	99US-0155659P.		
PR 28-SEP-1999;	99US-0156458P.		
PR 29-SEP-1999;	99US-0156596P.		
PR 04-OCT-1999;	99US-0157117P.		
PR 05-OCT-1999;	99US-0157753P.		
PR 06-OCT-1999;	99US-0157865P.		
PR 07-OCT-1999;	99US-0158029P.		
PR 08-OCT-1999;	99US-0158232P.		
PR 12-OCT-1999;	99US-0158369P.		
PR 13-OCT-1999;	99US-0159293P.		
PR 13-OCT-1999;	99US-0159294P.		
PR 13-OCT-1999;	99US-0159295P.		
PR 14-OCT-1999;	99US-0159329P.		
PR 14-OCT-1999;	99US-0159330P.		
PR 14-OCT-1999;	99US-0159331P.		
PR 14-OCT-1999;	99US-0159637P.		
PR 18-OCT-1999;	99US-0159638P.		
PR 21-OCT-1999;	99US-0160741P.		
PR 21-OCT-1999;	99US-0160767P.		
PR 21-OCT-1999;	99US-0160768P.		
PR 21-OCT-1999;	99US-0160770P.		
PR 21-OCT-1999;	99US-0160814P.		
PR 21-OCT-1999;	99US-0160815P.		
Query Match 43.0%; Score 656; DB 3; Length 303;			
Best Local Similarity 62.6%; Pred. No. 7.2e-60;			
Matches 127; Conservative 35; Mismatches 41; Indels 0; Gaps 0;			
QY 79 L L K R K S E V L F Y L N G R C I Y L V G M M G S G K S T V G K I M S E V L G Y S F F D S D K L V E Q A V G M P S V A 138			
DB 88 I L K R K A E V K P Y L N G R S M Y L V G M M G S G K T T V G K L M S K V L G Y T F F D C D T L I E Q M N G T S V A 147			
QY 139 Q I F K V H S E A F P R D N E S S V L R D L S S M R R L V A T G G A V I R P I N W R Y M K R G L S V M L D V P L D A 198			
DB 148 E I F V H H G E N F F R G K E T D A L K K L S S R Y Q V V S T G G A V I R P I N W K Y M H K G I S I W L D V P L E A 207			
QY 199 L A R R I A K Y G T A S R P L L D Q P S G D P Y A M A F S K L S M L A Q O R G D A Y A N A D V R S L E E I A K Q C H 258			
DB 208 L A H R I A A V G T D S R P L L H D E S G D A Y S V A F K R L S A I W D E R G E A Y T N A N A R V S L E N I A A K R G Y 267			
QY 259 D D V S K L T P T D I A E S L H K I E S F V 281			
DB 268 K N V S D L T P T E I C I E A F E Q V L S F L 290			
RESULT 3			
AAG50002			
ID	AAG50002 standard; protein; 220 AA.		
XX	AAG50002;		
XX			
DT	18-OCT-2000 (first entry)		
XX	Arabidopsis thaliana protein fragment SEQ ID NO: 63318.		
DE	Protein identification; signal transduction pathway; metabolic pathway;		
KW	hybridisation assay; genetic mapping; gene expression control; promoter;		
KW	termination sequence.		
XX	Arabidopsis thaliana.		
OS			
XX			
PN	EP1033405-A2.		
XX			
PD	06-SEP-2000.		
XX			
PF	25-FEB-2000; 2000EP-00301439.		
XX			
PR	25-FEB-1999; 99US-0121825P.		
PR	05-MAR-1999; 99US-0123180P.		
PR	09-MAR-1999; 99US-0123548P.		
PR	23-MAR-1999; 99US-0125788P.		
PR	25-MAR-1999; 99US-0126264P.		
PR	29-MAR-1999; 99US-0126785P.		
PR	01-APR-1999; 99US-0127462P.		
PR	06-APR-1999; 99US-0128234P.		
PR	08-APR-1999; 99US-0128714P.		
PR	16-APR-1999; 99US-0129845P.		
PR	19-APR-1999; 99US-0130077P.		
PR	21-APR-1999; 99US-0130449P.		
PR	23-APR-1999; 99US-0130510P.		
PR	28-APR-1999; 99US-0130891P.		
PR	30-APR-1999; 99US-0131449P.		
PR			

PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144086P.
PR 16-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144684P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144086P.
PR 16-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144684P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.

PR	21-OCT-1999;	99US-0160767P.	PR	21-APR-1999;	99US-0130449P.
PR	21-OCT-1999;	99US-0160768P.	PR	23-APR-1999;	99US-0130510P.
PR	21-OCT-1999;	99US-0160770P.	PR	23-APR-1999;	99US-0130891P.
PR	21-OCT-1999;	99US-0160814P.	PR	28-APR-1999;	99US-0131443P.
PR	21-OCT-1999;	99US-0160815P.	PR	30-APR-1999;	99US-0132048P.
PR	22-OCT-1999;	99US-0160980P.	PR	30-APR-1999;	99US-0132407P.
PR	22-OCT-1999;	99US-0160981P.	PR	04-MAY-1999;	99US-0132484P.
PR	22-OCT-1999;	99US-0160989P.	PR	05-MAY-1999;	99US-0132485P.
PR	25-OCT-1999;	99US-0161404P.	PR	06-MAY-1999;	99US-0132486P.
PR	25-OCT-1999;	99US-0161405P.	PR	06-MAY-1999;	99US-0132487P.
PR	25-OCT-1999;	99US-0161406P.	PR	07-MAY-1999;	99US-0132863P.
PR	26-OCT-1999;	99US-0161359P.	PR	11-MAY-1999;	99US-0134256P.
PR	26-OCT-1999;	99US-0161360P.	PR	14-MAY-1999;	99US-0134218P.
PR	26-OCT-1999;	99US-0161361P.	PR	14-MAY-1999;	99US-0134219P.
PR	28-OCT-1999;	99US-0161920P.	PR	14-MAY-1999;	99US-0134221P.
PR	28-OCT-1999;	99US-0161992P.	PR	14-MAY-1999;	99US-0134370P.
PR	28-OCT-1999;	99US-0161993P.	PR	18-MAY-1999;	99US-0134768P.
PR	29-OCT-1999;	99US-0162142P.	PR	19-MAY-1999;	99US-0134941P.
Query Match 42.6%; Score 650; DB 3; Length 220;			PR	20-MAY-1999;	99US-0135124P.
Best Local Similarity 62.7%; Pred. No. 1.9e-59;			PR	21-MAY-1999;	99US-0135353P.
Matches 126; Conservative 35; Mismatches 40; Indels 0; Gaps 0;			PR	24-MAY-1999;	99US-0135623P.
			PR	25-MAY-1999;	99US-0136021P.
QY	81	KRKSEEVLYLNGRCITLYLGMWGSKSTVGKIMSEVLGYSFFDSOKLVRQAVGMPSSVAQI 140	PR	27-MAY-1999;	99US-0136392P.
Db	7	QRKAEEVKPYLNGRSNYLYGMWGSKTTVTKLMSKVLGYTFDFCDTLIEQAMNGTSVAEI 66	PR	28-MAY-1999;	99US-0137822P.
			PR	01-JUN-1999;	99US-0137222P.
QY	141	FKVHSEAFPRDNESVLRDLSSNRRLVVTGGAVIRPINWRMKRGLSVWLVDVPLDALA 200	PR	03-JUN-1999;	99US-0137528P.
Db	67	FVHNGENFFRGKETDALKKLSSRYQVWSTGGGAVIRPINWKMKGISIWLDVPLEALA 126	PR	04-JUN-1999;	99US-0137502P.
QY	201	RIAKYGTASRPLLDQSPGDYPAMAFSKLSMLAQQRGDAYANADVRSLEETACKQGHDD 260	PR	07-JUN-1999;	99US-0137724P.
Db	127	HRIAAVGTDSRPLLDHSDGVSFAFKRUSAIWDERGEAYTNANARVSLNTAAKRGYKN 186	PR	08-JUN-1999;	99US-0138094P.
QY	261	VSKLTPTDIAIESLKHIESFV 281	PR	10-JUN-1999;	99US-0138540P.
Db	187	VSDLTPTETAIEAFEQVLSFL 207	PR	10-JUN-1999;	99US-0138847P.
			PR	14-JUN-1999;	99US-0139119P.
RESULT 4			PR	16-JUN-1999;	99US-0139452P.
ID	AAG50011 standard; protein; 220 AA.		PR	16-JUN-1999;	99US-0139453P.
XX			PR	17-JUN-1999;	99US-0139492P.
AC	AAG50011;		PR	18-JUN-1999;	99US-0139454P.
XX			PR	18-JUN-1999;	99US-0139455P.
DT	18-OCT-2000 (first entry)		PR	18-JUN-1999;	99US-0139456P.
DE	Arabidopsis thaliana protein fragment SEQ ID NO: 63330.		PR	18-JUN-1999;	99US-0139457P.
XX			PR	18-JUN-1999;	99US-0139458P.
XX			PR	18-JUN-1999;	99US-0139459P.
KW	Protein identification; signal transduction pathway; metabolic pathway;		PR	18-JUN-1999;	99US-0139460P.
KW	hybridisation assay; genetic mapping; gene expression control; promoter;		PR	18-JUN-1999;	99US-0139461P.
KW	termination sequence.		PR	18-JUN-1999;	99US-0139462P.
XX			PR	18-JUN-1999;	99US-0139463P.
OS	Arabidopsis thaliana.		PR	18-JUN-1999;	99US-0139750P.
XX			PR	18-JUN-1999;	99US-0139763P.
FN	EP1033405-A2.		PR	21-JUN-1999;	99US-0139817P.
PD	06-SEP-2000.		PR	21-JUN-1999;	99US-0139899P.
XX			PR	23-JUN-1999;	99US-0140353P.
XX			PR	23-JUN-1999;	99US-0140354P.
XX			PR	24-JUN-1999;	99US-0140695P.
XX			PR	28-JUN-1999;	99US-0140823P.
XX			PR	29-JUN-1999;	99US-0140991P.
XX			PR	30-JUN-1999;	99US-0141287P.
XX			PR	01-JUL-1999;	99US-0141842P.
XX			PR	01-JUL-1999;	99US-0142154P.
XX			PR	02-JUL-1999;	99US-0142055P.
XX			PR	06-JUL-1999;	99US-0142390P.
XX			PR	08-JUL-1999;	99US-0142803P.
XX			PR	09-JUL-1999;	99US-0142920P.
XX			PR	12-JUL-1999;	99US-0142977P.
XX			PR	13-JUL-1999;	99US-0143542P.
PR	25-FEB-1999;	99US-0121825P.	PR	14-JUL-1999;	99US-0143624P.
PR	05-MAR-1999;	99US-0123180P.	PR	15-JUL-1999;	99US-0144005P.
PR	09-MAR-1999;	99US-01233548P.	PR	16-JUL-1999;	99US-0144085P.
PR	23-MAR-1999;	99US-0125788P.	PR	16-JUL-1999;	99US-0144086P.
PR	25-MAR-1999;	99US-0126264P.	PR	19-JUL-1999;	99US-0144325P.
PR	29-MAR-1999;	99US-0126785P.	PR	19-JUL-1999;	99US-0144331P.
PR	01-APR-1999;	99US-0127462P.	PR	19-JUL-1999;	99US-0144332P.
PR	06-APR-1999;	99US-0128234P.	PR	19-JUL-1999;	99US-0144333P.
PR	08-APR-1999;	99US-0128714P.	PR	19-JUL-1999;	99US-0144334P.
PR	16-APR-1999;	99US-0129845P.	PR	19-JUL-1999;	99US-0144335P.
PR	19-APR-1999;	99US-0130077P.			

PR	20-JUL-1999;	99US-0144352P.
PR	20-JUL-1999;	99US-0144632P.
PR	20-JUL-1999;	99US-0144884P.
PR	21-JUL-1999;	99US-0144814P.
PR	21-JUL-1999;	99US-0145086P.
PR	21-JUL-1999;	99US-0145088P.
PR	21-JUL-1999;	99US-0145085P.
PR	22-JUL-1999;	99US-0145087P.
PR	22-JUL-1999;	99US-0145089P.
PR	22-JUL-1999;	99US-0145192P.
PR	22-JUL-1999;	99US-0145145P.
PR	23-JUL-1999;	99US-0145218P.
PR	23-JUL-1999;	99US-0145224P.
PR	26-JUL-1999;	99US-0145276P.
PR	27-JUL-1999;	99US-0145913P.
PR	27-JUL-1999;	99US-0145918P.
PR	27-JUL-1999;	99US-0145919P.
PR	28-JUL-1999;	99US-0145951P.
PR	02-AUG-1999;	99US-0146386P.
PR	02-AUG-1999;	99US-0146388P.
PR	02-AUG-1999;	99US-0146389P.
PR	03-AUG-1999;	99US-0147038P.
PR	04-AUG-1999;	99US-0147204P.
PR	04-AUG-1999;	99US-0147302P.
PR	05-AUG-1999;	99US-0147192P.
PR	05-AUG-1999;	99US-0147260P.
PR	06-AUG-1999;	99US-0147303P.
PR	06-AUG-1999;	99US-0147416P.
PR	06-AUG-1999;	99US-0147493P.
PR	09-AUG-1999;	99US-0147935P.
PR	10-AUG-1999;	99US-0148171P.
PR	11-AUG-1999;	99US-0148319P.
PR	12-AUG-1999;	99US-0148341P.
PR	13-AUG-1999;	99US-0148565P.
PR	13-AUG-1999;	99US-0148684P.
PR	16-AUG-1999;	99US-0149368P.
PR	17-AUG-1999;	99US-0149175P.
PR	18-AUG-1999;	99US-0149426P.
PR	20-AUG-1999;	99US-0149722P.
PR	20-AUG-1999;	99US-0149723P.
PR	20-AUG-1999;	99US-0149929P.
PR	23-AUG-1999;	99US-0149902P.
PR	23-AUG-1999;	99US-0149930P.
PR	25-AUG-1999;	99US-0150566P.
PR	26-AUG-1999;	99US-0150884P.
PR	27-AUG-1999;	99US-0151065P.
PR	27-AUG-1999;	99US-0151066P.
PR	27-AUG-1999;	99US-0151080P.
PR	30-AUG-1999;	99US-0151303P.
PR	31-AUG-1999;	99US-0151438P.
PR	01-SEP-1999;	99US-0151930P.
PR	07-SEP-1999;	99US-0152363P.
PR	10-SEP-1999;	99US-0153070P.
PR	13-SEP-1999;	99US-0153758P.
PR	15-SEP-1999;	99US-0154018P.
PR	16-SEP-1999;	99US-0154039P.
PR	20-SEP-1999;	99US-0154779P.
PR	23-SEP-1999;	99US-0155139P.
PR	23-SEP-1999;	99US-0155486P.
PR	24-SEP-1999;	99US-0155659P.
PR	28-SEP-1999;	99US-0155658P.
PR	29-SEP-1999;	99US-0156596P.
PR	04-OCT-1999;	99US-0157117P.
PR	05-OCT-1999;	99US-0157753P.
PR	06-OCT-1999;	99US-0157865P.
PR	07-OCT-1999;	99US-0158029P.
PR	08-OCT-1999;	99US-0158232P.
PR	12-OCT-1999;	99US-0158369P.
PR	13-OCT-1999;	99US-0159293P.
PR	13-OCT-1999;	99US-0159294P.
PR	13-OCT-1999;	99US-0159295P.
PR	14-OCT-1999;	99US-0159329P.
PR	14-OCT-1999;	99US-0159330P.
PR	14-OCT-1999;	99US-0159331P.
PR	14-OCT-1999;	99US-0159637P.
PR	14-OCT-1999;	99US-0159638P.
PR	18-OCT-1999;	99US-0159584P.
PR	21-OCT-1999;	99US-0160741P.
PR	21-OCT-1999;	99US-0160767P.
PR	21-OCT-1999;	99US-0160768P.
PR	21-OCT-1999;	99US-0160770P.
PR	21-OCT-1999;	99US-0160814P.
PR	21-OCT-1999;	99US-0160815P.
PR	22-OCT-1999;	99US-0160980P.
PR	22-OCT-1999;	99US-0160981P.
PR	22-OCT-1999;	99US-0160989P.
PR	25-OCT-1999;	99US-0161404P.
PR	25-OCT-1999;	99US-0161405P.
PR	25-OCT-1999;	99US-0161406P.
PR	26-OCT-1999;	99US-0161359P.
PR	26-OCT-1999;	99US-0161360P.
PR	26-OCT-1999;	99US-0161361P.
PR	28-OCT-1999;	99US-0161920P.
PR	28-OCT-1999;	99US-0161992P.
PR	28-OCT-1999;	99US-0161993P.
PR	29-OCT-1999;	99US-0162142P.
Query Match 42.6%; Score 650; DB 3; Length 220;		
Best Local Similarity 62.7%; Pred. No. 1.9e-59;		
Matches 126; Conservative 35; Mismatches 40; Indels 0; Gaps 0;		
QY	81	KKKSEVLFLNGRCIYLVGMGSGKSTVGKIMSEVLGVSPFDSKLVQAVGMPSVAOI 140
Db	7	QKRAEVPYLNGRSMYLVGMGSGKTTVGKLMKSVLGYTFDCDTLIEQAMNGTSVAEI 66
QY	141	FKVHSEAFPRDNESVLRDLSSMRLLVWATGGGAVIRPINRMRYMKRGLSVWLDVPLDALA 200
Db	67	FVHGHENFERGKETDALKKLSRYQVWVSTGGGAVIRPINRMRYMKHGISIWLDPLEALA 126
QY	201	RIIAKVGTSARPLLPQSGDPYAMAFSKLSMLAQOQGDYANADVRSLEEIAACKQCHDD 260
Db	127	HRIAAGTDSRPLLDHDESGDAYSVAFKRLSAIWDERGEAYTNANARVSLNIAAKRGYKN 186
QY	261	VSKLPTPTDIAIESLHKIESFV 281
Db	187	VSDLPTPTAIAEAFQVLSFL 207
RESULT 5		
AAG50001		
ID	AAG50001	standard; protein; 250 AA.
XX		
AC	AAG50001;	
XX		
DT	18-OCT-2000	(first entry)
XX		
DE	Arabidopsis thaliana	protein fragment SEQ ID NO: 63317.
XX		
KW	Protein identification; signal transduction pathway; metabolic pathway;	
KW	hybridisation assay; genetic mapping; gene expression control; promoter;	
KW	termination sequence.	
XX		
OS	Arabidopsis thaliana.	
XX		
PN	EP1033405-A2.	
XX		
XX	06-SEP-2000.	
XX		
PD	25-FEB-2000; 2000EP-00301439.	
XX		
XX	25-FEB-1999;	99US-0121825P.
PR	05-MAR-1999;	99US-0123180P.
PR	09-MAR-1999;	99US-0123548P.
PR	23-MAR-1999;	99US-0125788P.
PR	25-MAR-1999;	99US-0126264P.
PR	29-MAR-1999;	99US-0126785P.

PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140895P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.

PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150844P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 28-SEP-1999; 99US-0155659P.
PR 29-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.

PR	13-OCT-1999;	99US-0159293P.	PR	05-MAR-1999;	99US-01233180P.
PR	13-OCT-1999;	99US-0159294P.	PR	03-MAR-1999;	99US-01233548P.
PR	13-OCT-1999;	99US-0159295P.	PR	23-MAR-1999;	99US-0125788P.
PR	14-OCT-1999;	99US-0159329P.	PR	23-MAR-1999;	99US-0126264P.
PR	14-OCT-1999;	99US-0159330P.	PR	23-MAR-1999;	99US-0126785P.
PR	14-OCT-1999;	99US-0159331P.	PR	01-APR-1999;	99US-0127462P.
PR	14-OCT-1999;	99US-0159637P.	PR	08-APR-1999;	99US-0128234P.
PR	14-OCT-1999;	99US-0159638P.	PR	08-APR-1999;	99US-0128714P.
PR	18-OCT-1999;	99US-0159584P.	PR	16-APR-1999;	99US-0129845P.
PR	21-OCT-1999;	99US-0160741P.	PR	15-APR-1999;	99US-0130077P.
PR	21-OCT-1999;	99US-0160767P.	PR	21-APR-1999;	99US-0130449P.
PR	21-OCT-1999;	99US-0160770P.	PR	23-APR-1999;	99US-0130510P.
PR	21-OCT-1999;	99US-0160814P.	PR	23-APR-1999;	99US-0130891P.
PR	21-OCT-1999;	99US-0160815P.	PR	28-APR-1999;	99US-0131449P.
PR	22-OCT-1999;	99US-0160980P.	PR	30-APR-1999;	99US-0132048P.
PR	22-OCT-1999;	99US-0160981P.	PR	30-APR-1999;	99US-0132407P.
PR	22-OCT-1999;	99US-0160989P.	PR	04-MAY-1999;	99US-0132484P.
PR	25-OCT-1999;	99US-0161404P.	PR	05-MAY-1999;	99US-0132485P.
PR	25-OCT-1999;	99US-0161405P.	PR	06-MAY-1999;	99US-0132486P.
PR	25-OCT-1999;	99US-0161406P.	PR	06-MAY-1999;	99US-0132487P.
PR	26-OCT-1999;	99US-0161359P.	PR	07-MAY-1999;	99US-0132863P.
PR	26-OCT-1999;	99US-0161360P.	PR	11-MAY-1999;	99US-0134256P.
PR	26-OCT-1999;	99US-0161361P.	PR	14-MAY-1999;	99US-0134218P.
PR	28-OCT-1999;	99US-0161920P.	PR	14-MAY-1999;	99US-0134219P.
PR	28-OCT-1999;	99US-0161922P.	PR	14-MAY-1999;	99US-0134221P.
PR	28-OCT-1999;	99US-0161933P.	PR	14-MAY-1999;	99US-0134370P.
PR	29-OCT-1999;	99US-0162142P.	PR	18-MAY-1999;	99US-0134768P.
			PR	20-MAY-1999;	99US-0134941P.
			PR	20-MAY-1999;	99US-0135124P.
			PR	21-MAY-1999;	99US-0135353P.
			PR	24-MAY-1999;	99US-0135629P.
			PR	25-MAY-1999;	99US-0136021P.
			PR	27-MAY-1999;	99US-0136392P.
			PR	28-MAY-1999;	99US-0136782P.
Qy	81	KRKSEEVLYLNGRCIYLVGMGSGKSTVGKIMSEVLGYSPFSDSKLVEQAVGMPSVAQI	140	01-JUN-1999;	99US-0137222P.
Db	37	QRKAEEVKPYLNGRSMYLVGMGSGKTTVGKMSKVLGYTFDFCDTUEQAMNGTSVAEI	96	03-JUN-1999;	99US-0137528P.
			PR	04-JUN-1999;	99US-0137502P.
Qy	141	FKVHSEAFFRDNESVLRLDLSMERLVVATGGGAVIRPINRWYMKRGLSVWLDVPLDALA	200	07-JUN-1999;	99US-0137724P.
Db	97	FVHHGHNFFRGKETDALKCLSSRYQVVVSTGGGAVIRPINWYMKHKGISWLDVPLEALA	156	08-JUN-1999;	99US-0138094P.
			PR	10-JUN-1999;	99US-0138540P.
Qy	201	RRIAKVGTASRPLLDQSPDYAMAFSKLSMAQQRGDVANADVRSLEEIACKQGHDD	260	10-JUN-1999;	99US-0138847P.
Db	157	HRIAAVGTDSRPLLDHSDGDAYSVAFKRLSAIWDGERGEAYTNANARVSLENIAAKRGYKN	216	14-JUN-1999;	99US-0139119P.
			PR	16-JUN-1999;	99US-0139452P.
Qy	261	VSKLPTDPTDIAESLHKTESFV	281	16-JUN-1999;	99US-0139453P.
Db	217	VSDLTPTDTEIAEAFEQWLISFL	237	17-JUN-1999;	99US-0139492P.
			PR	18-JUN-1999;	99US-0139454P.
			PR	18-JUN-1999;	99US-0139455P.
			PR	18-JUN-1999;	99US-0139456P.
			PR	18-JUN-1999;	99US-0139457P.
			PR	18-JUN-1999;	99US-0139458P.
			PR	18-JUN-1999;	99US-0139459P.
			PR	18-JUN-1999;	99US-0139460P.
			PR	18-JUN-1999;	99US-0139461P.
			PR	18-JUN-1999;	99US-0139462P.
			PR	18-JUN-1999;	99US-0139463P.
			PR	18-JUN-1999;	99US-0139750P.
			PR	18-JUN-1999;	99US-0139763P.
			PR	21-JUN-1999;	99US-0139817P.
			PR	22-JUN-1999;	99US-0139899P.
			PR	23-JUN-1999;	99US-0140353P.
			PR	23-JUN-1999;	99US-0140354P.
			PR	24-JUN-1999;	99US-0140695P.
			PR	28-JUN-1999;	99US-0140823P.
			PR	29-JUN-1999;	99US-0140991P.
			PR	30-JUN-1999;	99US-0142887P.
			PR	01-JUL-1999;	99US-0141842P.
			PR	01-JUL-1999;	99US-0142154P.
			PR	02-JUL-1999;	99US-0142055P.
			PR	06-JUL-1999;	99US-0142390P.
			PR	08-JUL-1999;	99US-0142803P.
			PR	09-JUL-1999;	99US-0142920P.
			PR	12-JUL-1999;	99US-0142977P.
			PR	13-JUL-1999;	99US-0143542P.
			PR		
			PR		
			PR		
			PR		
			PR		
			PR		
			PR		
			PR		
			PR		
			PR		
			PR		
			PR		
			PR		
			PR		
			PR		
			PR		
			PR		
			PR		
			PR		
			PR		
			PR		
			PR		
			PR		
			PR		
			PR		
			PR		
			PR		
			PR		
			PR		
			PR		
			PR		
			PR		
			PR		
			PR		
			PR		
			PR		
			PR		
			PR		
			PR		
			PR		
			PR		
			PR		
			PR		
			PR		
			PR		
			PR		
			PR		
			PR		
			PR		
			PR		
			PR		
			PR		
			PR		
			PR		
			PR		
			PR		
			PR		
			PR		
			PR		
			PR		
			PR		
			PR		
			PR		
			PR		
			PR		
			PR		
			PR		
			PR		
			PR		
			PR		
			PR		
			PR		
			PR		
			PR		
			PR		
			PR		
			PR		
			PR		
			PR		
			PR		
			PR		
			PR		
			PR		
			PR		
			PR		
			PR		
			PR		
			PR		
			PR		
			PR		
			PR		
			PR		
			PR		
			PR		
			PR		
			PR		
			PR		
			PR		
			PR		
			PR		

PR 14-JUL-1999;	99US-0143624P.	PR 05-OCT-1999;	99US-0157753P.
PR 15-JUL-1999;	99US-0144005P.	PR 06-OCT-1999;	99US-0157865P.
PR 16-JUL-1999;	99US-0144085P.	PR 07-OCT-1999;	99US-0158023P.
PR 16-JUL-1999;	99US-0144086P.	PR 08-OCT-1999;	99US-0158232P.
PR 19-JUL-1999;	99US-0144325P.	PR 12-OCT-1999;	99US-0158369P.
PR 19-JUL-1999;	99US-0144331P.	PR 13-OCT-1999;	99US-0159293P.
PR 19-JUL-1999;	99US-0144332P.	PR 13-OCT-1999;	99US-0159294P.
PR 19-JUL-1999;	99US-0144333P.	PR 13-OCT-1999;	99US-0159295P.
PR 19-JUL-1999;	99US-0144334P.	PR 14-OCT-1999;	99US-0159323P.
PR 19-JUL-1999;	99US-0144335P.	PR 14-OCT-1999;	99US-0159330P.
PR 20-JUL-1999;	99US-0144352P.	PR 14-OCT-1999;	99US-0159331P.
PR 20-JUL-1999;	99US-0144632P.	PR 14-OCT-1999;	99US-0159637P.
PR 20-JUL-1999;	99US-0144884P.	PR 14-OCT-1999;	99US-0159638P.
PR 21-JUL-1999;	99US-0144814P.	PR 18-OCT-1999;	99US-0159584P.
PR 21-JUL-1999;	99US-0145086P.	PR 21-OCT-1999;	99US-0160741P.
PR 21-JUL-1999;	99US-0145088P.	PR 21-OCT-1999;	99US-0160767P.
PR 22-JUL-1999;	99US-0145085P.	PR 21-OCT-1999;	99US-0160768P.
PR 22-JUL-1999;	99US-0145087P.	PR 21-OCT-1999;	99US-0160770P.
PR 22-JUL-1999;	99US-0145089P.	PR 21-OCT-1999;	99US-0160814P.
PR 22-JUL-1999;	99US-0145192P.	PR 21-OCT-1999;	99US-0160815P.
PR 23-JUL-1999;	99US-0145145P.	PR 22-OCT-1999;	99US-0160980P.
PR 23-JUL-1999;	99US-0145218P.	PR 22-OCT-1999;	99US-0160981P.
PR 23-JUL-1999;	99US-0145224P.	PR 22-OCT-1999;	99US-0160989P.
PR 26-JUL-1999;	99US-0145276P.	PR 25-OCT-1999;	99US-0161404P.
PR 27-JUL-1999;	99US-0145913P.	PR 25-OCT-1999;	99US-0161405P.
PR 27-JUL-1999;	99US-0145918P.	PR 25-OCT-1999;	99US-0161406P.
PR 27-JUL-1999;	99US-0145919P.	PR 25-OCT-1999;	99US-0161359P.
PR 28-JUL-1999;	99US-0145951P.	PR 26-OCT-1999;	99US-0161360P.
PR 02-AUG-1999;	99US-0146386P.	PR 26-OCT-1999;	99US-0161361P.
PR 02-AUG-1999;	99US-0146388P.	PR 28-OCT-1999;	99US-0161920P.
PR 02-AUG-1999;	99US-0146389P.	PR 28-OCT-1999;	99US-0161992P.
PR 03-AUG-1999;	99US-0147038P.	PR 28-OCT-1999;	99US-0161993P.
PR 04-AUG-1999;	99US-0147204P.	PR 29-OCT-1999;	99US-0162142P.
PR 04-AUG-1999;	99US-0147302P.		
PR 05-AUG-1999;	99US-0147192P.		
PR 05-AUG-1999;	99US-0147260P.		
PR 06-AUG-1999;	99US-0147303P.		
PR 06-AUG-1999;	99US-0147416P.		
PR 09-AUG-1999;	99US-0147493P.		
PR 09-AUG-1999;	99US-0147935P.		
PR 10-AUG-1999;	99US-0148171P.		
PR 11-AUG-1999;	99US-0148319P.		
PR 12-AUG-1999;	99US-0148341P.		
PR 13-AUG-1999;	99US-0148565P.		
PR 13-AUG-1999;	99US-0148684P.		
PR 16-AUG-1999;	99US-0149368P.		
PR 17-AUG-1999;	99US-0149175P.		
PR 18-AUG-1999;	99US-0149426P.		
PR 20-AUG-1999;	99US-0149722P.		
PR 20-AUG-1999;	99US-0149723P.		
PR 20-AUG-1999;	99US-0149929P.		
PR 23-AUG-1999;	99US-0149902P.		
PR 23-AUG-1999;	99US-0149930P.		
PR 25-AUG-1999;	99US-0150566P.		
PR 26-AUG-1999;	99US-0150884P.		
PR 27-AUG-1999;	99US-0151065P.		
PR 27-AUG-1999;	99US-0151066P.		
PR 27-AUG-1999;	99US-0151080P.		
PR 30-AUG-1999;	99US-0151303P.		
PR 31-AUG-1999;	99US-0151438P.		
PR 01-SEP-1999;	99US-0151930P.		
PR 07-SEP-1999;	99US-0152363P.		
PR 10-SEP-1999;	99US-0153070P.		
PR 13-SEP-1999;	99US-0153758P.		
PR 15-SEP-1999;	99US-0154018P.		
PR 16-SEP-1999;	99US-0154039P.		
PR 20-SEP-1999;	99US-0154779P.		
PR 22-SEP-1999;	99US-0155139P.		
PR 23-SEP-1999;	99US-0155486P.		
PR 24-SEP-1999;	99US-0155659P.		
PR 28-SEP-1999;	99US-0156458P.		
PR 29-SEP-1999;	99US-0156596P.		
PR 04-OCT-1999;	99US-0157117P.		
<hr/>			
Query Match 42.6%; Score 650; DB 3; Length 250;			
Best Local Similarity 62.7%; Pred. No. 2.3e-59;			
Matches 126; Conservative 35; Mismatches 40; Indels 0; Gaps 0;			
QY 81	KKKEEVLFYINGRCIYLVGMGSGKSTVGKIMSEVLGYSFSDSKLYEQAVGMSVAQI	140	
DB 37	QRKAEVFPYLNGRSMYLVGMGSGKTTVGKLSKVLYGTFDFCDTLIEQAMNGTSVAEI	96	
QY 141	FKVHSEAFFRDNESSVLRLDLSMRRLVVATGGGAVIRPINRWYMKRGLSVMLDVLPLDALA	200	
DB 97	FVHGENFRGKETDALKLSRYQVVYSTGGGAVIRPINWKYMHKGISWLDVPLEALA	156	
QY 201	RIAKVGTASRPLLDQPSGDPYAMAFSKLSMLAQOQGDAYANADVRSLEETACKQHDD	260	
DB 157	HRIAAVGTDSPRLHDESGDAYSVAFKRLSAIWDERGEAYTNANARVSLNIAAKRGYKN	216	
QY 261	VSKLTPTDIAIESLHKIESFV	281	
DB 217	VSDLTPTEIAIEAFEQVLSFL	237	
RESULT 7			
ABB91833			
ID	ABB91833 standard; protein; 292 AA.		
XX	AC		
XX	ABB91833;		
XX	31-MAY-2002 (first entry)		
DT			
DE	Herbicideally active polypeptide SEQ ID NO 1044.		
XX	Herbicideal; plant; agriculture; herbicide.		
XX	Arabidopsis thaliana.		
OS			
XX	WO200210210-A2.		
PN			
XX	07-FEB-2002.		
PD			
XX			

PF 28-AUG-2001; 2001WO-EP009892.
XX 28-AUG-2001; 2001WO-EP009892.
XX (FARB) BAYER AG.
PI Tietjen K, Weidler M;
XX
DR WPI; 2002-269010/31.
XX
XX Identifying plant target proteins for herbicidally active compounds,
PT comprising aligning and comparing nucleic acid or amino acid sequences
PT from plant with nucleic acid or amino acid sequences from non-plant
PT organisms.
XX
XX Claim 5; SEQ ID NO 1044; 261pp + Sequence Listing; English.
XX
XX The invention relates to identifying target proteins (ABB90790-ABB94016)
CC for herbicidally active compounds, comprising aligning and comparing
CC nucleic acid or amino acid sequences from plant with nucleic acid or
CC amino acid sequences from non-plant organisms using suitable search
CC parameters, where plant sequences having an E-value greater by a factor
CC of 3 than the E-value of most similar non-plant sequences are selected.
CC The polypeptides or nucleic acids encoding them are useful for
CC identifying modulators. The identified modulators are useful as
CC herbicides
XX
SQ Sequence 292 AA;
Query Match 42.6%; Score 649.5; DB 5; Length 292;
Best Local Similarity 62.1%; Pred. No. 3.3e-59;
Matches 128; Conservative 36; Mismatches 41; Indels 1; Gaps 1;
QY 79 LKRRSEVLFLNGRCIYLVGMGSGKSTVGKIMSEVLGYSFFDSDKLVBOAVGMPVA 138
:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 88 ILKRAEVEKPYLNGRSMVLGVMGSGKTVGKLSKVLYGYFFDCDTLIEQAWNGTSVA 147
QY 139 QIPKVHSEAFFRDNESSVLRDLSSMRRLVATGGGAVIRPINRWYMKRGLSVWLDVPLDA 198
:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 148 EIFVHGENFFRGKETDALKLSSRYQVVVSTGGGAVIRPINWKYMHKGISIWLDPLEA 207
QY 199 LABRIAKVTASRPLADQSGDPYMAFSKLSMLAQOGRDAYANADVRSLEIACKQGH 258
:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 208 LAHRIAAGVTDSPRLHDSGGDAYSVAFKRLSAIWDERGEAYTNANARVSLNIAKRGY 267
:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 259 DVSKLTPTDIAIESLHKIESFVIEH 284
:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 268 KNVSDLTPTREIAIE-VSQVLSQLLHH 292
RESULT 8
AAG49994
ID AAG49994 standard; protein; 199 AA.
XX
XX AAG49994;
AC
XX
XX 18-OCT-2000 (first entry)
DT
DE Arabidopsis thaliana protein fragment SEQ ID NO: 63307.
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
XX Arabidopsis thaliana.
OS
XX
XX EP1033405-A2.
PN
XX
XX 06-SEP-2000.
PD
XX
XX 25-FEB-2000; 2000EP-00301439.
PF
XX
XX 25-FEB-1999; 99US-0121825P.
PR

PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0130891P.
PR 30-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 04-MAY-1999; 99US-0132407P.
PR 05-MAY-1999; 99US-0132484P.
PR 06-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 07-MAY-1999; 99US-0132487P.
PR 11-MAY-1999; 99US-0132863P.
PR 14-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 21-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.

PR	23-SEP-1999;	99US-0155486P.	XX	OS	Arabidopsis thaliana.
PR	24-SEP-1999;	99US-0155659P.	XX	OS	Arabidopsis thaliana.
PR	28-SEP-1999;	99US-0156458P.	XX	PN	EP1033405-A2.
PR	29-SEP-1999;	99US-0156596P.	XX	XX	
PR	04-OCT-1999;	99US-0157117P.	XX	PD	06-SEP-2000.
PR	05-OCT-1999;	99US-0157753P.	XX	XX	
PR	06-OCT-1999;	99US-0157865P.	XX	XX	25-FEB-2000; 2000EP-00301439.
PR	07-OCT-1999;	99US-0158029P.	XX	XX	
PR	08-OCT-1999;	99US-0158232P.	XX	XX	
PR	12-OCT-1999;	99US-0158369P.	PR	25-FEB-1999;	99US-0121825P.
PR	13-OCT-1999;	99US-0159293P.	PR	05-MAR-1999;	99US-0123180P.
PR	13-OCT-1999;	99US-0159294P.	PR	09-MAR-1999;	99US-0123548P.
PR	13-OCT-1999;	99US-0159295P.	PR	23-MAR-1999;	99US-0125788P.
PR	14-OCT-1999;	99US-0159329P.	PR	25-MAR-1999;	99US-0126264P.
PR	14-OCT-1999;	99US-0159330P.	PR	29-MAR-1999;	99US-0126785P.
PR	14-OCT-1999;	99US-0159331P.	PR	01-APR-1999;	99US-0127462P.
PR	14-OCT-1999;	99US-0159637P.	PR	06-APR-1999;	99US-0128234P.
PR	14-OCT-1999;	99US-0159638P.	PR	08-APR-1999;	99US-0128714P.
PR	18-OCT-1999;	99US-0159584P.	PR	16-APR-1999;	99US-0129845P.
PR	21-OCT-1999;	99US-0160741P.	PR	19-APR-1999;	99US-0130077P.
PR	21-OCT-1999;	99US-0160767P.	PR	21-APR-1999;	99US-0130449P.
PR	21-OCT-1999;	99US-0160768P.	PR	23-APR-1999;	99US-0130510P.
PR	21-OCT-1999;	99US-0160770P.	PR	23-APR-1999;	99US-0130891P.
PR	21-OCT-1999;	99US-0160814P.	PR	30-APR-1999;	99US-0131449P.
PR	21-OCT-1999;	99US-0160815P.	PR	30-APR-1999;	99US-0132048P.
PR	22-OCT-1999;	99US-0160980P.	PR	30-APR-1999;	99US-0132407P.
PR	22-OCT-1999;	99US-0160981P.	PR	04-MAY-1999;	99US-0132484P.
PR	22-OCT-1999;	99US-0160989P.	PR	05-MAY-1999;	99US-0132485P.
PR	25-OCT-1999;	99US-0161404P.	PR	06-MAY-1999;	99US-0132486P.
PR	25-OCT-1999;	99US-0161405P.	PR	06-MAY-1999;	99US-0132487P.
PR	26-OCT-1999;	99US-0161359P.	PR	07-MAY-1999;	99US-0132863P.
PR	26-OCT-1999;	99US-0161360P.	PR	11-MAY-1999;	99US-0134256P.
PR	26-OCT-1999;	99US-0161361P.	PR	14-MAY-1999;	99US-0134218P.
PR	26-OCT-1999;	99US-0161920P.	PR	14-MAY-1999;	99US-0134219P.
PR	28-OCT-1999;	99US-0161992P.	PR	14-MAY-1999;	99US-0134221P.
PR	28-OCT-1999;	99US-0161993P.	PR	14-MAY-1999;	99US-0134370P.
PR	28-OCT-1999;	99US-0161993P.	PR	18-MAY-1999;	99US-0134768P.
PR	29-OCT-1999;	99US-0162142P.	PR	19-MAY-1999;	99US-0134941P.
			PR	20-MAY-1999;	99US-0135124P.
			PR	21-MAY-1999;	99US-0135353P.
			PR	24-MAY-1999;	99US-0135629P.
			PR	25-MAY-1999;	99US-0136021P.
			PR	27-MAY-1999;	99US-0136392P.
			PR	28-MAY-1999;	99US-0136782P.
			PR	01-JUN-1999;	99US-0137222P.
			PR	03-JUN-1999;	99US-0137528P.
			PR	04-JUN-1999;	99US-0137502P.
			PR	07-JUN-1999;	99US-0137724P.
			PR	08-JUN-1999;	99US-0138094P.
			PR	10-JUN-1999;	99US-0138540P.
			PR	10-JUN-1999;	99US-0138847P.
			PR	14-JUN-1999;	99US-0139119P.
			PR	16-JUN-1999;	99US-0139452P.
			PR	16-JUN-1999;	99US-0139453P.
			PR	17-JUN-1999;	99US-0139492P.
			PR	18-JUN-1999;	99US-0139454P.
			PR	18-JUN-1999;	99US-0139455P.
			PR	18-JUN-1999;	99US-0139456P.
			PR	18-JUN-1999;	99US-0139457P.
			PR	18-JUN-1999;	99US-0139458P.
			PR	18-JUN-1999;	99US-0139459P.
			PR	18-JUN-1999;	99US-0139460P.
			PR	18-JUN-1999;	99US-0139461P.
			PR	18-JUN-1999;	99US-0139462P.
			PR	18-JUN-1999;	99US-0139463P.
			PR	18-JUN-1999;	99US-0139750P.
			PR	18-JUN-1999;	99US-0139763P.
			PR	21-JUN-1999;	99US-0139817P.
			PR	22-JUN-1999;	99US-0139899P.
			PR	23-JUN-1999;	99US-0140353P.
			PR	23-JUN-1999;	99US-0140354P.
			PR	24-JUN-1999;	99US-0140695P.
			PR	28-JUN-1999;	99US-0140823P.

Query Match 39.5%; Score 603, DB 3; Length 199;
Best Local Similarity 62.4%; Pred. No. 1.4e-54;
Matches 116; Conservative 33; Mismatches 37; Indels 0; Gaps 0;
QY 96 IYLVGMGSGKTVGKIMSEVLGYSFFDSKLVQAVGMPSVAQIFKVHSEAFPRDNESS 155
Db 1 MTLVGMGSGKTVGKLMSEKVLGYFFDCTLIEQWNGTSAEIVFVHGENFFRGKETD 60
QY 156 VLRLSSMRRLVATGGGAVIRPINWRYMKRGLSVLWDLPLDALARRIAKVGTA SRPLLD 215
Db 61 ALKLLSRVQVVSTGGAVIRPINWKYHKGISLWDLVPLEALAHRIA AVGTD SRPLIH 120
QY 216 QPSGDPYNAFSLKSLMIAQORDAYANADVRSLEBIACKQGHDDVSKLTPPTDIAIESLH 275
Db 121 DESGDAYSVAFKRLSAIWDERGEAYTNANARVSLENIAARGYKNVSDLTPTTEIAIEAPE 180
QY 276 KIESFV 281
Db 181 QVLSFL 186

RESULT 10
AAG50012
ID AAG50012 standard; protein; 199 AA.
XX AC
XX AAG50012;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 63331.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.

PR 29-JUN-1999;	99US-0140991P.	PR 13-SEP-1999;	99US-0153758P.
PR 30-JUN-1999;	99US-0141287P.	PR 15-SEP-1999;	99US-0154018P.
PR 01-JUL-1999;	99US-0141842P.	PR 16-SEP-1999;	99US-0154039P.
PR 01-JUL-1999;	99US-0142154P.	PR 20-SEP-1999;	99US-0154779P.
PR 02-JUL-1999;	99US-0142055P.	PR 22-SEP-1999;	99US-0155139P.
PR 06-JUL-1999;	99US-0142390P.	PR 23-SEP-1999;	99US-0155486P.
PR 08-JUL-1999;	99US-0142803P.	PR 24-SEP-1999;	99US-0155659P.
PR 09-JUL-1999;	99US-0142920P.	PR 28-SEP-1999;	99US-0156458P.
PR 12-JUL-1999;	99US-0142977P.	PR 29-SEP-1999;	99US-0156596P.
PR 13-JUL-1999;	99US-0143542P.	PR 04-OCT-1999;	99US-0157117P.
PR 14-JUL-1999;	99US-0143624P.	PR 05-OCT-1999;	99US-0157753P.
PR 15-JUL-1999;	99US-0144005P.	PR 06-OCT-1999;	99US-0157865P.
PR 16-JUL-1999;	99US-0144086P.	PR 07-OCT-1999;	99US-0158029P.
PR 16-JUL-1999;	99US-0144308P.	PR 08-OCT-1999;	99US-0158232P.
PR 19-JUL-1999;	99US-0144325P.	PR 12-OCT-1999;	99US-0158369P.
PR 19-JUL-1999;	99US-0144331P.	PR 13-OCT-1999;	99US-0159293P.
PR 19-JUL-1999;	99US-0144332P.	PR 13-OCT-1999;	99US-0159294P.
PR 19-JUL-1999;	99US-0144333P.	PR 13-OCT-1999;	99US-0159295P.
PR 19-JUL-1999;	99US-0144334P.	PR 14-OCT-1999;	99US-0159329P.
PR 19-JUL-1999;	99US-0144335P.	PR 14-OCT-1999;	99US-0159330P.
PR 20-JUL-1999;	99US-0144352P.	PR 14-OCT-1999;	99US-0159331P.
PR 20-JUL-1999;	99US-0144632P.	PR 14-OCT-1999;	99US-0159637P.
PR 20-JUL-1999;	99US-0144684P.	PR 14-OCT-1999;	99US-0159638P.
PR 21-JUL-1999;	99US-0144814P.	PR 18-OCT-1999;	99US-0159584P.
PR 21-JUL-1999;	99US-0145086P.	PR 21-OCT-1999;	99US-0160741P.
PR 21-JUL-1999;	99US-0145088P.	PR 21-OCT-1999;	99US-0160767P.
PR 22-JUL-1999;	99US-0145085P.	PR 21-OCT-1999;	99US-0160768P.
PR 22-JUL-1999;	99US-0145087P.	PR 21-OCT-1999;	99US-0160770P.
PR 22-JUL-1999;	99US-0145089P.	PR 21-OCT-1999;	99US-0160814P.
PR 22-JUL-1999;	99US-0145192P.	PR 21-OCT-1999;	99US-0160815P.
PR 23-JUL-1999;	99US-0145145P.	PR 22-OCT-1999;	99US-0160980P.
PR 23-JUL-1999;	99US-0145218P.	PR 22-OCT-1999;	99US-0160981P.
PR 23-JUL-1999;	99US-0145224P.	PR 22-OCT-1999;	99US-0160989P.
PR 26-JUL-1999;	99US-0145276P.	PR 25-OCT-1999;	99US-0161404P.
PR 27-JUL-1999;	99US-0145913P.	PR 25-OCT-1999;	99US-0161405P.
PR 27-JUL-1999;	99US-0145918P.	PR 25-OCT-1999;	99US-0161406P.
PR 27-JUL-1999;	99US-0145919P.	PR 26-OCT-1999;	99US-0161359P.
PR 28-JUL-1999;	99US-0145951P.	PR 26-OCT-1999;	99US-0161360P.
PR 02-AUG-1999;	99US-0146386P.	PR 26-OCT-1999;	99US-0161361P.
PR 02-AUG-1999;	99US-0146388P.	PR 28-OCT-1999;	99US-0161920P.
PR 02-AUG-1999;	99US-0146389P.	PR 28-OCT-1999;	99US-0161922P.
PR 03-AUG-1999;	99US-0147038P.	PR 28-OCT-1999;	99US-0161993P.
PR 04-AUG-1999;	99US-0147204P.	PR 29-OCT-1999;	99US-0162142P.
PR 04-AUG-1999;	99US-0147302P.		
PR 05-AUG-1999;	99US-0147192P.		
PR 05-AUG-1999;	99US-0147260P.		
PR 06-AUG-1999;	99US-0147303P.		
PR 06-AUG-1999;	99US-0147416P.		
PR 09-AUG-1999;	99US-0147493P.		
PR 09-AUG-1999;	99US-0147935P.		
PR 10-AUG-1999;	99US-0148171P.		
PR 11-AUG-1999;	99US-0148319P.		
PR 12-AUG-1999;	99US-0148341P.		
PR 13-AUG-1999;	99US-0148565P.		
PR 13-AUG-1999;	99US-0148684P.		
PR 16-AUG-1999;	99US-0149368P.		
PR 17-AUG-1999;	99US-0149175P.		
PR 18-AUG-1999;	99US-0149426P.		
PR 20-AUG-1999;	99US-0149722P.		
PR 20-AUG-1999;	99US-0149723P.		
PR 20-AUG-1999;	99US-0149929P.		
PR 23-AUG-1999;	99US-0149902P.		
PR 23-AUG-1999;	99US-0149930P.		
PR 25-AUG-1999;	99US-0150566P.		
PR 26-AUG-1999;	99US-0150884P.		
PR 27-AUG-1999;	99US-0151065P.		
PR 27-AUG-1999;	99US-0151066P.		
PR 27-AUG-1999;	99US-0151080P.		
PR 30-AUG-1999;	99US-0151303P.		
PR 31-AUG-1999;	99US-0151438P.		
PR 01-SEP-1999;	99US-0151930P.		
PR 07-SEP-1999;	99US-0152363P.		
PR 10-SEP-1999;	99US-0153070P.		
Query Match 39.5%; Score 603; DB 3; Length 199;		Best Local Similarity 62.4%; Pred. No. 1.4e-54;	
Matches 116; Conservative 33; Mismatches 37; Indels 0; Gaps 0;			
QY	96 IYLVGMGSGKSTVGKIMSEVLGYSFDFSDKLVEQAVGMPVQAQIFKVHSEAFRDNESS 155		
Db	1 MYLVGMGSGKTTVGKLSKVLGYTFFDCDTLIEQAMNGTSVAEIFVHHGENFPRGKETD 60		
QY	156 VLRLDSSNRRLVATGGGAVIRPINRWYMKRGLSVWLDPDLALARRIAKYGTASRPILD 215		
Db	61 ALKKLSSRYQVVSTGGGAVIRPINRWYMKHGISIWLDPLEALAHRIAAGVTSRPLH 120		
QY	216 QPSGDPYAMAFSKLSMLAQQRDAYANADVRVSEELIEACKQCHDDVSKLTPDTDAIESLH 275		
Db	121 DESGDAYSVAFKRLSAIWDERGEAYTNANARVLENIAAKRGYKNVSDLTPTTEIAIEAFE 180		
QY	276 KIESFV 281		
Db	181 QVLSFL 186		
RESULT 11			
AAG42178			
ID	AAG42178 standard; protein; 301 AA.		
XX			
AC	AAG42178;		
XX			
DT	18-OCT-2000 (first entry)		
XX			

DE Arabidopsis thaliana protein fragment SEQ ID NO: 52568.
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-00301439.
XX
XX 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132486P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134376P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 21-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.

PR	30-AUG-1999;	99US-0151303P.
PR	31-AUG-1999;	99US-0151438P.
PR	01-SEP-1999;	99US-0151930P.
PR	07-SEP-1999;	99US-0152363P.
PR	10-SEP-1999;	99US-0153070P.
PR	13-SEP-1999;	99US-0153758P.
PR	15-SEP-1999;	99US-0154018P.
PR	16-SEP-1999;	99US-0154039P.
PR	20-SEP-1999;	99US-0154779P.
PR	22-SEP-1999;	99US-0155139P.
PR	23-SEP-1999;	99US-0155486P.
PR	24-SEP-1999;	99US-0155659P.
PR	28-SEP-1999;	99US-0156458P.
PR	29-SEP-1999;	99US-0156596P.
PR	04-OCT-1999;	99US-0157117P.
PR	05-OCT-1999;	99US-0157753P.
PR	06-OCT-1999;	99US-0157865P.
PR	07-OCT-1999;	99US-0158029P.
PR	08-OCT-1999;	99US-0158232P.
PR	12-OCT-1999;	99US-0158369P.
PR	13-OCT-1999;	99US-0159293P.
PR	13-OCT-1999;	99US-0159294P.
PR	13-OCT-1999;	99US-0159295P.
PR	14-OCT-1999;	99US-0159329P.
PR	14-OCT-1999;	99US-0159330P.
PR	14-OCT-1999;	99US-0159331P.
PR	14-OCT-1999;	99US-0159637P.
PR	18-OCT-1999;	99US-0159638P.
PR	18-OCT-1999;	99US-0159584P.
PR	21-OCT-1999;	99US-0160741P.
PR	21-OCT-1999;	99US-0160767P.
PR	21-OCT-1999;	99US-0160768P.
PR	21-OCT-1999;	99US-0160770P.
PR	21-OCT-1999;	99US-0160814P.
PR	21-OCT-1999;	99US-0160815P.
PR	22-OCT-1999;	99US-0160980P.
PR	22-OCT-1999;	99US-0160981P.
PR	22-OCT-1999;	99US-0160989P.
PR	25-OCT-1999;	99US-0161404P.
PR	25-OCT-1999;	99US-0161405P.
PR	25-OCT-1999;	99US-0161406P.
PR	26-OCT-1999;	99US-0161359P.
PR	26-OCT-1999;	99US-0161360P.
PR	26-OCT-1999;	99US-0161361P.
PR	28-OCT-1999;	99US-0161920P.
PR	28-OCT-1999;	99US-0161922P.
PR	28-OCT-1999;	99US-0161933P.
PR	29-OCT-1999;	99US-0162142P.
Query Match 39.4%; Score 602; DB 3; Length 301;		
Best Local Similarity 54.9%; Pred. No. 3.4e-54;		
Matches 129; Conservative 38; Mismatches 58; Indels 10; Gaps 5;		
QY	57 RLRA---KSSGGHENS---HNSVDE--ALLLRKSEEVLYLNGRCIYLVGMGSGKST 108	
Db	55 RLRSVSDKNSALLTCGLSLHSPDEEQILKQKAAEVKPYLNGRSMYLVGMGSGKTT 114	
QY	109 VGKIMSEVLGYSPFDSKLVQAGVMPVAQIFKVHSEAFFRDNESVLRDLSSM-RLV 167	
Db	115 VGKIMARSIGYTFDFCDTLIEQAMKGTVAEIFHFGESVREKETEALKKLSMYHQV 174	
QY	168 VATGGGAVIRPINWYMKRGLSVWLDVPLDALARRIAKVGTASRPILL-DQPSGDPYAMAF 226	
Db	175 VSTGGGAVIRPINWYMKRGLSVWLDVPLDALARRIAKVGTASRPILLHDDESGDTYTAAL 234	
QY	227 SKLSMLAAQQRDAYANADVRSLEETACKQGHDDVSKLFTPTDIAIESLHKIESFV 281	
Db	235 NRLSTIWDARGEAYTKASARVSLNITLKLGRVSRVSDLTPABIAIEAFQVQSYL 289	

RESULT 12
AAG42177
ID AAG42177 standard; protein; 316 AA.

XX	AAG42177;	
AC	18-OCT-2000 (first entry)	
XX		
DT		
XX		
DE	Arabidopsis thaliana protein fragment SEQ ID NO: 52567.	
XX		
KW	Protein identification; signal transduction pathway; metabolic pathway;	
KW	hybridisation assay; genetic mapping; gene expression control; promoter;	
KW	termination sequence.	
XX		
OS	Arabidopsis thaliana.	
XX		
PN	EP1033405-A2.	
XX		
PD	06-SEP-2000.	
XX		
PF	25-FEB-2000; 2000EP-00301439.	
XX		
PR	25-FEB-1999; 99US-0121825P.	
PR	05-MAR-1999; 99US-0123180P.	
PR	09-MAR-1999; 99US-0123548P.	
PR	23-MAR-1999; 99US-0125788P.	
PR	25-MAR-1999; 99US-0126264P.	
PR	29-MAR-1999; 99US-0126785P.	
PR	01-APR-1999; 99US-0127462P.	
PR	06-APR-1999; 99US-0128234P.	
PR	08-APR-1999; 99US-0128714P.	
PR	16-APR-1999; 99US-0129845P.	
PR	19-APR-1999; 99US-0130077P.	
PR	21-APR-1999; 99US-0130449P.	
PR	23-APR-1999; 99US-0130510P.	
PR	23-APR-1999; 99US-0130891P.	
PR	28-APR-1999; 99US-0131449P.	
PR	30-APR-1999; 99US-0132048P.	
PR	30-APR-1999; 99US-0132407P.	
PR	04-MAY-1999; 99US-0132484P.	
PR	05-MAY-1999; 99US-0132485P.	
PR	06-MAY-1999; 99US-0132486P.	
PR	06-MAY-1999; 99US-0132487P.	
PR	07-MAY-1999; 99US-0132863P.	
PR	11-MAY-1999; 99US-0134256P.	
PR	14-MAY-1999; 99US-0134218P.	
PR	14-MAY-1999; 99US-0134219P.	
PR	14-MAY-1999; 99US-0134221P.	
PR	14-MAY-1999; 99US-0134370P.	
PR	18-MAY-1999; 99US-0134768P.	
PR	19-MAY-1999; 99US-0134941P.	
PR	20-MAY-1999; 99US-0135124P.	
PR	21-MAY-1999; 99US-0135353P.	
PR	24-MAY-1999; 99US-0135629P.	
PR	25-MAY-1999; 99US-0136021P.	
PR	27-MAY-1999; 99US-0136392P.	
PR	28-MAY-1999; 99US-0136782P.	
PR	01-JUN-1999; 99US-0137222P.	
PR	03-JUN-1999; 99US-0137528P.	
PR	04-JUN-1999; 99US-0137502P.	
PR	07-JUN-1999; 99US-0137724P.	
PR	08-JUN-1999; 99US-0138094P.	
PR	10-JUN-1999; 99US-0138540P.	
PR	10-JUN-1999; 99US-0138847P.	
PR	14-JUN-1999; 99US-0139119P.	
PR	16-JUN-1999; 99US-0139452P.	
PR	16-JUN-1999; 99US-0139453P.	
PR	17-JUN-1999; 99US-0139492P.	
PR	18-JUN-1999; 99US-0139454P.	
PR	18-JUN-1999; 99US-0139455P.	
PR	18-JUN-1999; 99US-0139456P.	
PR	18-JUN-1999; 99US-0139457P.	
PR	18-JUN-1999; 99US-0139458P.	
PR	18-JUN-1999; 99US-0139459P.	
PR	18-JUN-1999; 99US-0139460P.	
PR	18-JUN-1999; 99US-0139461P.	

RESULT 13
AAG12964
ID AAG12964 standard; protein; 199 AA.
XX
AC AAG12964;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 12280.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
hybridisation assay; genetic mapping; gene expression control; promoter;
termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0130891P.
PR 30-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 04-MAY-1999; 99US-0132407P.
PR 05-MAY-1999; 99US-0132484P.
PR 06-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 07-MAY-1999; 99US-0132487P.
PR 11-MAY-1999; 99US-0132863P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134376P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137528P.
PR 07-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.


```
Db      61  SRRYQVVSTGGGAVIRPNWKYMHKGISWLDVPLEALAHRIAAGVTDSPILLHDESGD 120
      || :.:||:|||||:|||||:| :.:||:|||||:|||||:| :|| :|||
Qy      221 PYAMAFSKLSMLAQQRGDAYANADVRSLEETACKQGHDDVSKLPTPTDIAIESLHKIESF 280
      ||:||||:| :.:||:|||||:|||||:| :.:||:|||||:|||||:| :|
Db      121 AYSVAFKRLSAIWDERGEAYTNANARVSLENIAAKRGYKNVSDLTPTETAIABFAFEQVLISF 180
      ||:||||:| :.:||:|||||:|||||:| :.:||:|||||:|||||:| :|
Qy      281 V 281
      |
Db      181 L 181

RESULT 15
AAG12965
ID  AAG12965 standard; protein; 194 AA.
XX
AC  AAG12965;
XX
XX
DT  17-OCT-2000 (first entry)
DE  Arabidopsis thaliana protein fragment SEQ ID NO: 12281.
XX
KW  Protein identification; signal transduction pathway; metabolic pathway;
KW  hybridisation assay; genetic mapping; gene expression control; promoter;
KW  termination sequence.
XX
XX  Arabidopsis thaliana.
OS
XX  EP1033405-A2.
XX
XX  06-SEP-2000.
XX
XX  25-FEB-2000; 2000EP-00301439.
XX
XX  25-FEB-1999; 99US-0121825P.
XX  05-MAR-1999; 99US-0123180P.
XX  09-MAR-1999; 99US-0123548P.
XX  23-MAR-1999; 99US-0125788P.
XX  25-MAR-1999; 99US-0126264P.
XX  29-MAR-1999; 99US-0126785P.
XX  01-APR-1999; 99US-0127462P.
XX  06-APR-1999; 99US-0128234P.
XX  08-APR-1999; 99US-0128714P.
XX  16-APR-1999; 99US-0129845P.
XX  19-APR-1999; 99US-0130077P.
XX  21-APR-1999; 99US-0130449P.
XX  23-APR-1999; 99US-0130510P.
XX  28-APR-1999; 99US-0130891P.
XX  30-APR-1999; 99US-0132048P.
XX  30-APR-1999; 99US-0132407P.
XX  04-MAY-1999; 99US-0132484P.
XX  05-MAY-1999; 99US-0132485P.
XX  06-MAY-1999; 99US-0132486P.
XX  06-MAY-1999; 99US-0132487P.
XX  07-MAY-1999; 99US-0132863P.
XX  11-MAY-1999; 99US-0134256P.
XX  14-MAY-1999; 99US-0134218P.
XX  14-MAY-1999; 99US-0134219P.
XX  14-MAY-1999; 99US-0134221P.
XX  14-MAY-1999; 99US-0134370P.
XX  18-MAY-1999; 99US-0134768P.
XX  19-MAY-1999; 99US-0134941P.
XX  20-MAY-1999; 99US-0135124P.
XX  21-MAY-1999; 99US-0135353P.
XX  24-MAY-1999; 99US-0135629P.
XX  25-MAY-1999; 99US-0136021P.
XX  27-MAY-1999; 99US-0136392P.
XX  28-MAY-1999; 99US-0136782P.
XX  01-JUN-1999; 99US-0137222P.
XX  03-JUN-1999; 99US-0137528P.
XX  04-JUN-1999; 99US-0137502P.
XX  07-JUN-1999; 99US-0137724P.
PR  08-JUN-1999; 99US-0138094P.
PR  10-JUN-1999; 99US-0138540P.
PR  10-JUN-1999; 99US-0138847P.
PR  14-JUN-1999; 99US-0139119P.
PR  16-JUN-1999; 99US-0139452P.
PR  16-JUN-1999; 99US-0139453P.
PR  17-JUN-1999; 99US-0139492P.
PR  18-JUN-1999; 99US-0139454P.
PR  18-JUN-1999; 99US-0139455P.
PR  18-JUN-1999; 99US-0139456P.
PR  18-JUN-1999; 99US-0139457P.
PR  18-JUN-1999; 99US-0139458P.
PR  18-JUN-1999; 99US-0139459P.
PR  18-JUN-1999; 99US-0139460P.
PR  18-JUN-1999; 99US-0139461P.
PR  18-JUN-1999; 99US-0139462P.
PR  18-JUN-1999; 99US-0139463P.
PR  18-JUN-1999; 99US-0139750P.
PR  18-JUN-1999; 99US-0139763P.
PR  21-JUN-1999; 99US-0139817P.
PR  22-JUN-1999; 99US-0139899P.
PR  23-JUN-1999; 99US-0140353P.
PR  23-JUN-1999; 99US-0140354P.
PR  24-JUN-1999; 99US-0140695P.
PR  28-JUN-1999; 99US-0140823P.
PR  29-JUN-1999; 99US-0140991P.
PR  30-JUN-1999; 99US-0141287P.
PR  01-JUL-1999; 99US-0141842P.
PR  01-JUL-1999; 99US-0142154P.
PR  02-JUL-1999; 99US-0142055P.
PR  06-JUL-1999; 99US-0142390P.
PR  08-JUL-1999; 99US-0142803P.
PR  09-JUL-1999; 99US-0142920P.
PR  12-JUL-1999; 99US-0142977P.
PR  13-JUL-1999; 99US-0143542P.
PR  14-JUL-1999; 99US-0143624P.
PR  15-JUL-1999; 99US-0144005P.
PR  16-JUL-1999; 99US-0144085P.
PR  16-JUL-1999; 99US-0144086P.
PR  19-JUL-1999; 99US-0144325P.
PR  19-JUL-1999; 99US-0144331P.
PR  19-JUL-1999; 99US-0144332P.
PR  19-JUL-1999; 99US-0144333P.
PR  19-JUL-1999; 99US-0144334P.
PR  19-JUL-1999; 99US-0144335P.
PR  20-JUL-1999; 99US-0144352P.
PR  20-JUL-1999; 99US-0144632P.
PR  20-JUL-1999; 99US-0144884P.
PR  21-JUL-1999; 99US-0144814P.
PR  21-JUL-1999; 99US-0145086P.
PR  21-JUL-1999; 99US-0145088P.
PR  22-JUL-1999; 99US-0145089P.
PR  22-JUL-1999; 99US-0145087P.
PR  22-JUL-1999; 99US-0145089P.
PR  22-JUL-1999; 99US-0145192P.
PR  23-JUL-1999; 99US-0145145P.
PR  23-JUL-1999; 99US-0145218P.
PR  23-JUL-1999; 99US-0145224P.
PR  26-JUL-1999; 99US-0145276P.
PR  27-JUL-1999; 99US-0145913P.
PR  27-JUL-1999; 99US-0145918P.
PR  27-JUL-1999; 99US-0145919P.
PR  28-JUL-1999; 99US-0145951P.
PR  02-AUG-1999; 99US-0146386P.
PR  02-AUG-1999; 99US-0146388P.
PR  02-AUG-1999; 99US-0146389P.
PR  03-AUG-1999; 99US-0147038P.
PR  04-AUG-1999; 99US-0147204P.
PR  04-AUG-1999; 99US-0147302P.
PR  05-AUG-1999; 99US-0147192P.
PR  05-AUG-1999; 99US-0147260P.
PR  06-AUG-1999; 99US-0147303P.
PR  06-AUG-1999; 99US-0147416P.
```

Qy	101	MMGSGKSTVGKIMSEVLCYSPFSDSDKLVEQAVGHPSVAQIKVUISEAFPRNESSVLRLD	160
Db	1	MMGSGKTTVGKLMGKVLGYTFDFCDTLIEQAMNGTSAEIVFVHGENFVRGKETDALKL	60
Qy	161	SSMERLVVATGGGAVIRPINWRYMKRGISVWLVDPLDALARRIAKVGTSAPRLLDQPSD	220
Db	61	SSRYQVVVSTGGGAVIRPINWRYMHKGISVWLDVPELAHRIAAVGTDSRPLLDHESGD	120
Qy	221	PYAMAFSKLSMLAQGRDAYANADVRSLEBIAKQGHDDVSKLTPDTDAIESLHKIESF	280
Db	121	AYSVAFKLSAIWDERGEAYTNANARVSNLENIAAKGYKNVSDLTPTTEICIEAFQVLSF	180
Qy	281	V 281	
Db	181	L 181	

Search completed: August 25, 2005, 11:17:05
Job time : 170 secs

PR	09-AUG-1999;	99US-0147493P.
PR	09-AUG-1999;	99US-0147935P.
PR	10-AUG-1999;	99US-0148171P.
PR	11-AUG-1999;	99US-0148319P.
PR	12-AUG-1999;	99US-0148341P.
PR	13-AUG-1999;	99US-0148565P.
PR	13-AUG-1999;	99US-0148684P.
PR	16-AUG-1999;	99US-0149368P.
PR	17-AUG-1999;	99US-0149175P.
PR	18-AUG-1999;	99US-0149426P.
PR	20-AUG-1999;	99US-0149722P.
PR	20-AUG-1999;	99US-0149723P.
PR	20-AUG-1999;	99US-0149929P.
PR	23-AUG-1999;	99US-0149902P.
PR	23-AUG-1999;	99US-0149930P.
PR	25-AUG-1999;	99US-0150566P.
PR	26-AUG-1999;	99US-0150884P.
PR	27-AUG-1999;	99US-0151065P.
PR	27-AUG-1999;	99US-0151066P.
PR	27-AUG-1999;	99US-0151080P.
PR	30-AUG-1999;	99US-0151303P.
PR	31-AUG-1999;	99US-0151438P.
PR	01-SEP-1999;	99US-0151930P.
PR	07-SEP-1999;	99US-0152363P.
PR	10-SEP-1999;	99US-0153070P.
PR	13-SEP-1999;	99US-0153758P.
PR	15-SEP-1999;	99US-0154018P.
PR	16-SEP-1999;	99US-0154039P.
PR	20-SEP-1999;	99US-0154779P.
PR	22-SEP-1999;	99US-0155139P.
PR	23-SEP-1999;	99US-0155486P.
PR	24-SEP-1999;	99US-0155659P.
PR	28-SEP-1999;	99US-0156458P.
PR	29-SEP-1999;	99US-0156596P.
PR	04-OCT-1999;	99US-0157117P.
PR	05-OCT-1999;	99US-0157753P.
PR	06-OCT-1999;	99US-0157865P.
PR	07-OCT-1999;	99US-0158029P.
PR	08-OCT-1999;	99US-0158232P.
PR	12-OCT-1999;	99US-0158369P.
PR	13-OCT-1999;	99US-0159293P.
PR	13-OCT-1999;	99US-0159294P.
PR	13-OCT-1999;	99US-0159295P.
PR	14-OCT-1999;	99US-0159329P.
PR	14-OCT-1999;	99US-0159330P.
PR	14-OCT-1999;	99US-0159331P.
PR	14-OCT-1999;	99US-0159637P.
PR	14-OCT-1999;	99US-0159638P.
PR	18-OCT-1999;	99US-0159584P.
PR	21-OCT-1999;	99US-0160741P.
PR	21-OCT-1999;	99US-0160767P.
PR	21-OCT-1999;	99US-0160768P.
PR	21-OCT-1999;	99US-0160770P.
PR	21-OCT-1999;	99US-0160814P.
PR	21-OCT-1999;	99US-0160815P.
PR	22-OCT-1999;	99US-0160980P.
PR	22-OCT-1999;	99US-0160981P.
PR	25-OCT-1999;	99US-0160989P.
PR	25-OCT-1999;	99US-0161404P.
PR	25-OCT-1999;	99US-0161405P.
PR	25-OCT-1999;	99US-0161406P.
PR	26-OCT-1999;	99US-0161359P.
PR	26-OCT-1999;	99US-0161360P.
PR	26-OCT-1999;	99US-0161361P.
PR	28-OCT-1999;	99US-0161920P.
PR	28-OCT-1999;	99US-0161992P.
PR	28-OCT-1999;	99US-0161993P.
PR	29-OCT-1999;	99US-0162142P.

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 25, 2005, 10:58:14 ; Search time 22 Seconds
(without alignments)
1034.907 Million cell updates/sec

Title: US-10-660-226-10
Perfect score: 1526
Sequence: 1 MEAGGVGLALQARAAGFGSS.....ADSSASDAQESQIQRIQTL 305
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues
Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/prodata/1/iaa/5A COMB.pep:*
2: /cgn2_6/prodata/1/iaa/5B COMB.pep:*
3: /cgn2_6/prodata/1/iaa/6A COMB.pep:*
4: /cgn2_6/prodata/1/iaa/6B COMB.pep:*
5: /cgn2_6/prodata/1/iaa/PTUS COMB.pep:*
6: /cgn2_6/prodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	243.5	16.0	231	4	US-09-540-236-2087
2	222	14.5	200	4	US-09-328-352-6149
3	222	14.5	200	4	US-09-489-039A-10812
4	212.5	13.9	191	4	US-09-543-681A-6682
5	205.5	13.5	184	4	US-09-252-991A-32964
6	185.5	12.2	178	4	US-09-107-532A-5800
7	179	11.7	158	4	US-09-583-110-3547
8	179	11.7	159	4	US-09-107-433-3398
9	167.5	11.0	170	4	US-09-134-000C-4979
10	162	10.6	177	4	US-09-543-681A-5313
11	145	9.5	213	4	US-09-489-039A-13538
12	143.5	9.4	184	4	US-09-198-452A-1113
13	143.5	9.4	184	4	US-09-438-185A-1039
14	140.5	9.2	172	3	US-09-134-001C-3610
15	125	8.2	276	4	US-09-248-796A-18195
16	125	8.2	1551	3	US-09-425-665-2
17	125	8.2	1551	3	US-09-685-668-2
18	103	6.7	1049	4	US-09-252-991A-25336
19	100	6.6	789	3	US-08-727-308-1
20	97	6.4	515	4	US-09-902-540-14617
21	96	6.3	467	4	US-09-252-991A-28862
22	95.5	6.3	353	4	US-09-252-991A-26103
23	95	6.2	2459	3	US-09-413-814-70
24	95	6.2	2439	3	US-09-335-409-7
25	95	6.2	2439	3	US-09-568-102-7
26	95	6.2	2439	3	US-09-567-969-7
27	95	6.2	2439	3	US-09-568-480-7

28	95	6.2	2439	3	US-09-568-486-7	Sequence 7, Appli
29	95	6.2	2439	3	US-09-568-472-7	Sequence 7, Appli
30	95	6.2	2439	3	US-09-567-899-7	Sequence 7, Appli
31	94	6.2	698	4	US-09-828-062-9	Sequence 9, Appli
32	94	6.2	1832	3	US-09-335-409-4	Sequence 4, Appli
33	94	6.2	1832	3	US-09-568-102-4	Sequence 4, Appli
34	94	6.2	1832	3	US-09-567-969-4	Sequence 4, Appli
35	94	6.2	1832	3	US-09-568-480-4	Sequence 4, Appli
36	94	6.2	1832	3	US-09-568-486-4	Sequence 4, Appli
37	94	6.2	1832	3	US-09-568-472-4	Sequence 4, Appli
38	94	6.2	1832	3	US-09-567-893-4	Sequence 4, Appli
39	93.5	6.1	171	4	US-09-328-352-8227	Sequence 8227, Ap
40	93.5	6.1	616	4	US-09-949-016-6889	Sequence 6889, Ap
41	93.5	6.1	616	4	US-09-949-016-7740	Sequence 7740, Ap
42	93	6.1	1413	4	US-09-252-991A-23627	Sequence 23627, A
43	92.5	6.1	549	4	US-09-902-540-10128	Sequence 10128, A
44	92.5	6.1	947	2	US-08-500-857A-4	Sequence 4, Appli
45	92	6.0	209	2	US-08-793-229-29	Sequence 29, Appli

ALIGNMENTS

RESULT 1
US-09-540-236-2087
; Sequence 2087, Application US/09540236
; Patent No. 6673910
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATA
; FILE REFERENCE: 2709.2005-001
; CURRENT APPLICATION NUMBER: US/09/540,236
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 3840
; SEQ ID NO 2087
; LENGTH: 231
; TYPE: PRT
; ORGANISM: M.catarrhalis
US-09-540-236-2087

Query Match 16.0%; Score 243.5; DB 4; Length 231;
Best Local Similarity 30.9%; Pred. No. 1.3e-17;
Matches 69; Conservative 44; Mismatches 75; Indels 35; Gaps 8;

QY 85 EVLVFYLNGR-----CIVLVGMWGSGKSTVGKIMSEVLGYSPFSDSKLVQAVGMPVS 137
DB 30 ESWVVEKAGNALSKOLPAIFLVGPNAGKTTIGKLAKHLGRNFIDCDWYIYDQTG-ADI 88
QY 138 AQIFKVHSEAFPRDNSESVLRDLSSMRLLVWATGGVIRPINWYMKRGLSVWLDVPLD 197
DB 89 PWIFEKEGEGFRDRETRALQELTALPMIATGGAGVGRTKNRELLKGLVIYLDASVD 148
QY 198 ALARIKAVGTASRPLLDPQSGDPYAMAFSKLSMLAQQRGDYANADVRSLEEIAACKQG 257
DB 149 TOLARTKK--DKNRPLLQ--AQNPRAV---LESYKKRHPLY-----REVA----- 187
QY 258 HDDVSKLT----PTDIAIESLHKIESFVIEHTADSSASDAQAE 296
DB 188 --DIILLTGRAYPKOMIGEILLEVENYALDHAPNKLKTDQHAH 228

RESULT 2
US-09-328-352-6149
; Sequence 6149, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04

```
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 6682
; LENGTH: 191
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-6682

Query Match      13.9%; Score 212.5; DB 4; Length 191;
Best Local Similarity 34.2%; Pred. No. 2e-14;
Matches 64; Conservative 30; Mismatches 70; Indels 23; Gaps 7;

QY   94 RCIVLVGMGSGKSTVGKTMSEVLGVSPFDSKLVEQAVCMPSVAQIFKVHSEAPRDNE 153
     |||:|||||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db    23 RNIFLVPGMGAGKTIGRLAQQLSMEFYSDQEIERGTG-ADVGVWFDVEGEGRQRE 81

QY   154 SSVLRDLSSMRRLVATGGAV-IPIPNRYMKRGLSVWLDPDLDALARRIAKVGTSRP 212
     ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db    82 EKIINELTEKGQIVLATGGGSVKSRTRNRLSARGVVVYLETNIEKQLARTOR--DKKP 139

QY   213 LLD--QPSGDPPAMAFSKLSMLAQQRGDAYANADVRVSLEEIACKQG-DDVSKLTPTDI 269
     ||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db    140 LLQVEPVD-----VLETLAEENPLY-----EEIADIHTDDQSAKIVANQ 183

QY   270 AIESLHK 276
     ||| |
Db    184 IIELLEK 190

RESULT 5
US-09-252-991A-32964
; Sequence 32964, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 32964
; LENGTH: 184
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-32964

Query Match      13.5%; Score 205.5; DB 4; Length 184;
Best Local Similarity 32.8%; Pred. No. 1e-13;
Matches 63; Conservative 33; Mismatches 61; Indels 35; Gaps 7;

QY   66 GHEN----SHNSVDEALLKKRKSEEVLYINGRCIYLVGMMGSGKSTVGKIMSEVLGYSP 121
     :|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db    2  GHNRSLISNNVN-----LILVGPWGAGKSTIGRLAKELHLAF 41

QY   122 FDSDKLVEQAVCMPSVAQIFKVHSEAFFRDNESSVLRDLSSMRRLVATGGGAVIRPINW 181
     |||:|||||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db    42 KDSDBKEIQRCG-ANIPWFIVDFEGVFREFREQAMLTELCAADGMVIATGGGAVMRDGNR 100

QY   182 RYMKR-GLSVWLDPDALARRIKAVGTASRPLLDQSPGDPPYAMAFSKLSMLAQQRDAY 240
     ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db    101 QVLRAGRVVYLHASVEHQIARTAR--DRNRPLLQKNPCOI-----LEDLMALRDPly 152

QY   241 AN-ADRVSLSE 251
     ||| |
Db    153 REIADVVTETD 164

RESULT 6
```



```
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...159
; SEQUENCE DESCRIPTION: SEQ ID NO: 3398:
US-09-107-433-3398

Query Match      11.7%; Score 179; DB 4; Length 159;
Best Local Similarity 34.8%; Pred. No. 5.6e-11;
Matches 62; Conservative 27; Mismatches 63; Indels 26; Gaps 8;

QY 98 IYLVGMGSGKSTVGKIMSEVLGYSPFSDKLVQAVGMPVSAQIFKVHSEAFPRDNESVL 157
Db 1 IYLVGMGSGKSTVGKIMSEVLGYSPFSDKLVQAVGMPVSAQIFKVHSEAFPRDNESVL 157
6 LLGFMGAGKSTIAR-----GLDNYLMDALIEKRLGM-SIANFFAEKGBETPRQVESEVL 60

QY 158 RDLSSMRRLVATGGGAVIRPINRWYMKRGL-SVWLDVPLDALARRIAKVGTASRPILLDQ 216
Db 61 ADLLQTDQ-VVSTGGGVVISQRNRLTKNTNINYLKADFTLYQRIAADKDNQRPLFLN 119

QY 217 PGSDPYAMAFSKLSMLAQQRDAYANADVRSLEETACKQGHDDVSKLTPPTDIAIESL 274
Db 120 NSKE-----ELVAIFQERQAWY-----EEVASRV--LDVTKLSPEEI-IEEL 158

RESULT 9
US-09-134-000C-4979
; Sequence 4979, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4979
; LENGTH: 170
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-134-000C-4979

Query Match      11.0%; Score 167.5; DB 4; Length 170;
Best Local Similarity 28.3%; Pred. No. 1e-09;
Matches 52; Conservative 38; Mismatches 73; Indels 21; Gaps 6;

QY 96 IYLVGMGSGKSTVGKIMSEVLGYSPFSDKLVQAVGMPVSAQIFKVHSEAFPRDNES 155
Db 6 IYLVGMGAGKTIGQSLANKLMPHLDLTALIEKIGR-SIPDYFEKYGEAFAFREQETQ 64

QY 156 VLRLDSSMRRLVATGGGAVIRPINRWYMKRGLSV-WLDVPLDALARRIAKVGTASRPILL 214
Db 65 LKLELSK-NTAVLSTGGGIWVGPENRSLKSFQQVIYLHATPEELKRLITEDTENQRPLA 123

QY 215 DQPSGDPYAMAFSKLSMLAQQRDAYANADVRSLEETACKQGHDDVSKLTPPTDIAIESL 274
Db 124 IERS-----KEIITLFSRKNFY-----EE--CAKMTDTTNRSPPEEIIINEIL 165

QY 275 HKIE 278
Db 166 QQLK 169

RESULT 10
US-09-543-681A-5313
; Sequence 5313, Application US/09543681A
; Patent No. 6605709
```

```
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 5313
; LENGTH: 177
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-5313

Query Match      10.6%; Score 162; DB 4; Length 177;
Best Local Similarity 35.5%; Pred. No. 4.2e-09;
Matches 43; Conservative 22; Mismatches 52; Indels 4; Gaps 4;

QY 96 IYLVGMGSGKSTVGKIMSEVLGYSPFSDKLVQAVGMPVSAQIFKVHSEAFPRDNES 155
Db 12 IYLVGMGAGKTIGKALSLSLDYQFIDTNDWITDKY-QOTISSMVEAEGMDVFRQLSESD 70

QY 156 VLRLDSSMRRLVATGGGAVIRPINRWYMKR-GLSVWLDVPLDALARRIAKVGTASRPIL 213
Db 71 ALIQSQPNQ-VISTGGGIVLAERKRVYMKNSGVIVYLQASLETTLVERLSQDPNEAQRPS 129

QY 214 L 214
Db 130 L 130

RESULT 11
US-09-489-039A-13538
; Sequence 13538, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 13538
; LENGTH: 213
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-13538

Query Match      9.5%; Score 145; DB 4; Length 213;
Best Local Similarity 25.7%; Pred. No. 3.6e-07;
Matches 53; Conservative 39; Mismatches 66; Indels 48; Gaps 8;

QY 96 IYLVGMGSGKSTVGKIMSEVLGYSPFSDKLVQAVGMPVSAQIFKVHSEAFPRDNES 155
Db 41 IFLIGPRGCGKTTVGHALARARHFQSDTDHRL-QAHEQRTVAEIVQAEQWARFRELETL 99

QY 156 VLRLDSSMRRLVATGGGAVIRPINRWYMKR-RGLSVWLDVPLDALARRIAKVGTASRPIL 213
Db 100 LKAV-VLTPNTVIATGGGIVLAERQRMRENGVVILYQASVSALIDRLLEATPKAEQRPT 158

QY 214 LDQPSGDPYAMAFSKLSMLAQQRDAYANADVRSLEETACKQGHDDVSKLTPPTDIAIES 273
Db 159 L---TGKP-----VREEVGEVLAQR-----EA 177

QY 274 LHKIESFVIEHTADSSADAQAESQI 299
Db 178 LYRDAA---HHIVDATASPDVRVEQI 200
```

```
RESULT 12
US-09-198-452A-1113
; Sequence 1113, Application US/09198452A
; Patent No. 6559294
; GENERAL INFORMATION:
; APPLICANT: Griffais, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments thereof and uses thereof, in particular for the diagnosis, prevention and treatment of infection
; TITLE OF INVENTION: and treatment of infection
; FILE REFERENCE: 9710-003-999
; CURRENT APPLICATION NUMBER: US/09/198,452A
; CURRENT FILING DATE: 1998-11-24
; NUMBER OF SEQ ID NOS: 6849
; SEQ ID NO 1113
; LENGTH: 184
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-09-198-452A-1113

Query Match          9.4%; Score 143.5; DB 4; Length 184;
Best Local Similarity 26.5%; Pred. No. 4.2e-07;
Matches 56; Conservative 36; Mismatches 74; Indels 45; Gaps 8;

QY 92 NGRCYLVGMGSGKSTVGKIMSEVLGYGFFDSKLV---EQAVGMPVAQIFKVHSEAF 148
Db 7 NVMTIILCGLPTSGKSLGKALAKFLNLPFYDLDLIVSNYSALYSSAEIYKAYGDKQ 66
QY 149 FRDNESVLRDLSSMRRLVATGGAVIRPINWRMYK-RGLSVWLDVPLDALARRIAKV 207
Db 67 FSECEARILETLPPEDAL-ISLGGGTLMEASYRAIQTRGALVFLSVLPLIYERLEKRG 125
QY 208 TASRPLLDQPSGDPYAMAFSKLSMLAQQRGDAYANADVRSLEEIAKQGHDDVSKLTPT 267
Db 126 LPER--LKE-----AMTKPLSEILTERID-----RMKEIA-----DYI----- 157
QY 268 DIAESLHKIESFVIEHTADSSASDAQESQ 298
Db 158 -----FPVDHVDHSSKSSLEQASQ 176

RESULT 13
US-09-438-185A-1039
; Sequence 1039, Application US/09438185A
; Patent No. 6822071
; GENERAL INFORMATION:
; APPLICANT: Stephens, Richard
; APPLICANT: Mitchell, Wayne
; APPLICANT: Kalman, Sue
; APPLICANT: Davis, Ronald
; TITLE OF INVENTION: The Regents of the University of California
; TITLE OF INVENTION: Chlamydia pneumoniae Genome Sequence
; FILE REFERENCE: 018941-000411US
; CURRENT APPLICATION NUMBER: US/09/438,185A
; CURRENT FILING DATE: 2002-03-13
; PRIOR APPLICATION NUMBER: US 60/108,279
; PRIOR FILING DATE: 1998-11-12
; PRIOR APPLICATION NUMBER: US 60/128,606
; PRIOR FILING DATE: 1999-04-08
; NUMBER OF SEQ ID NOS: 1074
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1039
; LENGTH: 184
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
; FEATURE:
; OTHER INFORMATION: Cpn1038
US-09-438-185A-1039

Query Match          9.4%; Score 143.5; DB 4; Length 184;
Best Local Similarity 26.5%; Pred. No. 4.2e-07;
Matches 56; Conservative 36; Mismatches 74; Indels 45; Gaps 8;

QY 92 NGRCYLVGMGSGKSTVGKIMSEVLGYGFFDSKLV---EQAVGMPVAQIFKVHSEAF 148
Db 7 NVMTIILCGLPTSGKSLGKALAKFLNLPFYDLDLIVSNYSALYSSAEIYKAYGDKQ 66
QY 149 FRDNESVLRDLSSMRRLVATGGAVIRPINWRMYK-RGLSVWLDVPLDALARRIAKV 207
Db 67 FSECEARILETLPPEDAL-ISLGGGTLMEASYRAIQTRGALVFLSVLPLIYERLEKRG 125
QY 208 TASRPLLDQPSGDPYAMAFSKLSMLAQQRGDAYANADVRSLEEIAKQGHDDVSKLTPT 267
Db 126 LPER--LKE-----AMTKPLSEILTERID-----RMKEIA-----DYI----- 157
QY 268 DIAESLHKIESFVIEHTADSSASDAQESQ 298
Db 158 -----FPVDHVDHSSKSSLEQASQ 176
```

```
QY 92 NGRCYLVGMGSGKSTVGKIMSEVLGYGFFDSKLV---EQAVGMPVAQIFKVHSEAF 148
Db 7 NVMTIILCGLPTSGKSLGKALAKFLNLPFYDLDLIVSNYSALYSSAEIYKAYGDKQ 66
QY 149 FRDNESVLRDLSSMRRLVATGGAVIRPINWRMYK-RGLSVWLDVPLDALARRIAKV 207
Db 67 FSECEARILETLPPEDAL-ISLGGGTLMEASYRAIQTRGALVFLSVLPLIYERLEKRG 125
QY 208 TASRPLLDQPSGDPYAMAFSKLSMLAQQRGDAYANADVRSLEEIAKQGHDDVSKLTPT 267
Db 126 LPER--LKE-----AMTKPLSEILTERID-----RMKEIA-----DYI----- 157
QY 268 DIAESLHKIESFVIEHTADSSASDAQESQ 298
Db 158 -----FPVDHVDHSSKSSLEQASQ 176

RESULT 14
US-09-134-001C-3610
; Sequence 3610, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 3610
; LENGTH: 172
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3610

Query Match          9.2%; Score 140.5; DB 3; Length 172;
Best Local Similarity 28.9%; Pred. No. 7.9e-07;
Matches 46; Conservative 31; Mismatches 73; Indels 9; Gaps 6;

QY 96 IYLVGMGSGKSTVGKIMSEVLGYGFFDSKLVQAVGMPVAQIFKVHSEAFPRDNSS 155
Db 10 IILVGMFGTKTTVGKYLSDLNLSYVDLDNFIE-VNECKSPNIFNDIGEGFRSLR 68
QY 156 VLRLDSSMRRL-VVATGGAVIRPINWRMYKGLS-VMLDVPDLALARRIAKVGTASRPL 213
Db 69 YLK--SCLNTFDIISTGGGIIEDTNSLKLKKNQKHVWMLDCCDIEIFKRVK--NDSHRPN 124
QY 214 LDQPSGDPYAMAFSKLSMLAQQRGDAYANADVRSLEEI 252
Db 125 AKSKNQLDLDALYS--SRLSYNEIAFMKVDQAQSVSEI 161

RESULT 15
US-09-248-796A-18195
; Sequence 18195, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICA
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 18195
```

```
; LENGTH: 276
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-18195

Query Match      8.2%; Score 125; DB 4; Length 276;
Best Local Similarity 26.4%; Pred. No. 7.2e-05;
Matches 33; Conservative 28; Mismatches 44; Indels 20; Gaps 4;

QY 56 LRLRAKKSSGCH----NSHNSVDEALLLKKKSEEVLFYLNRCIYLVGNMGSKGSTVGKI 112
Db 72 LHSKFKEIDGYPPTDTHVDK-----SDKSIIVIGMRGTGKSTLSEW 117

QY 113 MSEVLGYSFFDSKLVQAVGMPVAQIFKVHSEAFFRDNESSVLRDLSS--MRRLAVAT 170
Db 118 LASFLGFKMLDMDKYLEEKLG-TGIKSLIKAKGWYFROBEAIVAKCEFTKFSKGYVLST 176

QY 171 GGGAV 175
Db 177 GGGIV 181
```

Search completed: August 25, 2005, 11:17:53
Job time : 23 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 25, 2005, 11:12:05 ; Search time 65 Seconds
(without alignments)
1843.731 Million cell updates/sec

Title: US-10-660-226-10
Perfect score: 1526
Sequence: 1 MEAGGVGLALQARAAGFGSS.....ADSSASDAQESQIRIQTL 305

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1767149 seqs, 392926209 residues
Total number of hits satisfying chosen parameters: 1767149

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US10E_PUBCOMB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
19: /cgn2_6/ptodata/2/pubpaa/US11A_PUBCOMB.pep.*
20: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
21: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
22: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1526	100.0	305	15	US-10-660-226-10
2	1526	100.0	314	15	Sequence 10, Appl
3	1457	95.5	303	16	Sequence 69542, A
4	1387	90.9	303	16	Sequence 201135,
5	1387	90.9	324	15	Sequence 201137,
6	1387	90.9	326	15	Sequence 62604, A
7	1351.5	88.6	330	16	Sequence 57834, A
8	1301.5	85.3	308	16	Sequence 44984, A
9	1159	76.0	273	15	Sequence 147803
10	1061	69.5	217	16	Sequence 28, Appl
11	886	58.1	311	16	Sequence 201136,
					Sequence 187425,

12	874	57.3	282	15	US-10-660-226-26	Sequence 26, Appl
13	860	56.4	245	15	US-10-660-226-20	Sequence 20, Appl
14	858	56.2	292	16	US-10-425-115-253534	Sequence 251534, A
15	841	55.1	231	15	US-10-660-226-12	Sequence 12, Appl
16	733	48.0	199	16	US-10-739-930-10076	Sequence 10076, A
17	700	45.9	259	16	US-10-437-963-104823	Sequence 104823, A
18	625.5	41.0	299	15	US-10-424-599-212900	Sequence 212900, A
19	625.5	41.0	301	15	US-10-425-114-54946	Sequence 54946, A
20	620	40.6	132	16	US-10-437-963-187427	Sequence 187427, A
21	606.5	39.7	300	16	US-10-739-930-6067	Sequence 6067, Ap
22	605	39.6	154	16	US-10-425-115-201132	Sequence 201132, A
23	578.5	37.9	155	15	US-10-660-226-16	Sequence 16, Appl
24	547	35.8	120	15	US-10-425-114-52290	Sequence 52290, A
25	535	35.1	170	15	US-10-424-599-280141	Sequence 280141, A
26	535	35.1	170	15	US-10-425-114-54844	Sequence 54844, A
27	532.5	34.9	231	16	US-10-437-963-104800	Sequence 104800, A
28	504	33.0	152	16	US-10-767-701-42830	Sequence 42830, A
29	423.5	27.8	220	15	US-10-424-599-212902	Sequence 212902, A
30	409	26.8	163	15	US-10-424-599-168580	Sequence 168580, A
31	405	26.5	95	15	US-10-424-599-205496	Sequence 205496, A
32	374	24.5	103	11	US-03-864-408A-8602	Sequence 8602, Ap
33	364.5	23.9	278	15	US-10-660-226-22	Sequence 22, Appl
34	345	22.6	278	15	US-10-424-599-154207	Sequence 154207, A
35	342	22.4	89	16	US-10-425-115-336215	Sequence 336215, A
36	333	21.8	135	15	US-10-424-599-224958	Sequence 224958, A
37	329.5	21.6	280	15	US-10-369-493-2620	Sequence 2620, Ap
38	318.5	20.9	189	15	US-10-369-493-19953	Sequence 19953, A
39	316.5	20.7	188	15	US-10-369-493-20994	Sequence 20994, A
40	307	20.1	176	15	US-10-424-599-277475	Sequence 277475, A
41	296	19.4	176	15	US-10-369-493-19110	Sequence 19110, A
42	294.5	19.3	134	15	US-10-660-226-14	Sequence 14, Appl
43	273	17.9	153	15	US-10-767-701-39076	Sequence 39076, A
44	270.5	17.7	138	16	US-10-660-226-14	Sequence 14, Appl
45	270.5	17.7	272	15	US-10-660-226-14	Sequence 14, Appl

ALIGNMENTS

RESULT 1
US-10-660-226-10
; Sequence 10, Application US/10660226
; Publication No. US2004006484A1
; GENERAL INFORMATION:
; APPLICANT: Cahoon, Rebecca E.
; APPLICANT: Falco, S. Carl
; APPLICANT: Farnodu, Layo O.
; APPLICANT: Hitz, William D.
; APPLICANT: Rendina, Alan
; TITLE OF INVENTION: Chorismate Biosynthesis Enzymes
; FILE REFERENCE: BB-1159-C
; CURRENT APPLICATION NUMBER: US/10/660,226
; CURRENT FILING DATE: 2003-09-11
; PRIOR APPLICATION NUMBER: US/09/354,501
; PRIOR FILING DATE: 1999-07-16
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/093,611
; PRIOR FILING DATE: EARLIER FILING DATE: July 21, 1998
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 10
; LENGTH: 305
; TYPE: PRT
; ORGANISM: Zea mays
US-10-660-226-10

Query Match 100.0%; Score 1526; DB 15; Length 305;
Best Local Similarity 100.0%; Pred. No. 6.8e-141;
Matches 305; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MEAGGVGLALQARAAGFGSSRRHGGLOAQTGSLRVADPAGPAVAVRARGSKVPAPLRLRA 60
DB 1 MEAGGVGLALQARAAGFGSSRRHGGLOAQTGSLRVADPAGPAVAVRARGSKVPAPLRLRA 60

```

QY 61 KKSOGGSHNSVDEALLKRSSEVLFYNGRCIYLVGMGSGKSTVGKIMSEVLGYS 120
Db 61 KKSOGGSHNSVDEALLKRSSEVLFYNGRCIYLVGMGSGKSTVGKIMSEVLGYS 120

QY 121 FFDSDKLVEQAVGMPSPAQIFKVHSEAFPRDNESVLRDLSSMRRLVVAATGGGAVIRPIN 180
Db 121 FFDSDKLVEQAVGMPSPAQIFKVHSEAFPRDNESVLRDLSSMRRLVVAATGGGAVIRPIN 180

QY 181 WRYMKRGLSVLWDVPLDALARRIAKVGTASRPDLDPSPGDPYAMAFSKLSMLAQQRGDY 240
Db 181 WRYMKRGLSVLWDVPLDALARRIAKVGTASRPDLDPSPGDPYAMAFSKLSMLAQQRGDY 240

QY 241 ANADVRVSLSEETACKQGHDDVSKLTPTDIAIESLHKIESFVIEHTADSSASDAQAESQIQ 300
Db 241 ANADVRVSLSEETACKQGHDDVSKLTPTDIAIESLHKIESFVIEHTADSSASDAQAESQIQ 300

QY 301 RIQTL 305
Db 301 RIQTL 305

RESULT 2
US-10-425-114-69542
; Sequence 69542, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 69542
; LENGTH: 314
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMFLB73062D01_FLI.pep
US-10-425-114-69542

```

```

Query Match 100.0%; Score 1526; DB 15; Length 314;
Best Local Similarity 100.0%; Pred. No. 7.1e-141;
Matches 305; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEAGGVGLALQARAAGFGSSRRHGGLOAPTGSURVADPAGPAVAVRARGSKPVAPLRLRA 60
Db 10 MEAGGVGLALQARAAGFGSSRRHGGLOAPTGSURVADPAGPAVAVRARGSKPVAPLRLRA 69

QY 61 KKSOGGSHNSVDEALLKRSSEVLFYNGRCIYLVGMGSGKSTVGKIMSEVLGYS 120
Db 70 KKSOGGSHNSVDEALLKRSSEVLFYNGRCIYLVGMGSGKSTVGKIMSEVLGYS 129

QY 121 FFDSDKLVEQAVGMPSPAQIFKVHSEAFPRDNESVLRDLSSMRRLVVAATGGGAVIRPIN 180
Db 130 FFDSDKLVEQAVGMPSPAQIFKVHSEAFPRDNESVLRDLSSMRRLVVAATGGGAVIRPIN 189

QY 181 WRYMKRGLSVLWDVPLDALARRIAKVGTASRPDLDPSPGDPYAMAFSKLSMLAQQRGDY 240
Db 190 WRYMKRGLSVLWDVPLDALARRIAKVGTASRPDLDPSPGDPYAMAFSKLSMLAQQRGDY 249

QY 241 ANADVRVSLSEETACKQGHDDVSKLTPTDIAIESLHKIESFVIEHTADSSASDAQAESQIQ 300
Db 250 ANADVRVSLSEETACKQGHDDVSKLTPTDIAIESLHKIESFVIEHTADSSASDAQAESQIQ 309

QY 301 RIQTL 305
Db 310 RIQTL 314

```

```

RESULT 3
US-10-425-115-201135
; Sequence 201135, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 201135
; LENGTH: 303
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_115020C.1.pep
US-10-425-115-201135

Query Match 95.5%; Score 1457; DB 16; Length 303;
Best Local Similarity 96.1%; Pred. No. 4e-134;
Matches 293; Conservative 2; Mismatches 8; Indels 2; Gaps 1;

QY 1 MEAGGVGLALQARAAGFGSSRRHGGLOAPTGSURVADPAGPAVAVRARGSKPVAPLRLRA 60
Db 1 MEAGGVGLALQARAAGFGSSRRHGGLOAPTGSURVADPAGPAVAVRARGSKPVAPLRLRA 58

QY 61 KKSOGGSHNSVDEALLKRSSEVLFYNGRCIYLVGMGSGKSTVGKIMSEVLGYS 120
Db 59 KKSOGGSHNSVDEALLKRSSEVLFYNGRCIYLVGMGSGKSTVGKIMSEVLGYS 118

QY 121 FFDSDKLVEQAVGMPSPAQIFKVHSEAFPRDNESVLRDLSSMRRLVVAATGGGAVIRPIN 180
Db 119 FFDSDKLVEQAVGMPSPAQIFKVHSEAFPRDNESVLRDLSSMRRLVVAATGGGAVIRPIN 178

QY 181 WRYMKRGLSVLWDVPLDALARRIAKVGTASRPDLDPSPGDPYAMAFSKLSMLAQQRGDY 240
Db 179 WRYMKRGLSVLWDVPLDALARRIAKVGTASRPDLDPSPGDPYAMAFSKLSMLAQQRGDY 238

QY 241 ANADVRVSLSEETACKQGHDDVSKLTPTDIAIESLHKIESFVIEHTADSSASDAQAESQIQ 300
Db 239 ANADVRVSLSEETACKQGHDDVSKLTPTDIAIESLHKIESFVIEHTADSSASDAQAESQIQ 298

QY 301 RIQTL 305
Db 299 RIQTL 303

RESULT 4
US-10-425-115-201137
; Sequence 201137, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 201137
; LENGTH: 303
; TYPE: PRT
; ORGANISM: Zea mays

```


FEATURE:
OTHER INFORMATION: Clone ID: MR74577_115022C.1.pap
US-10-425-115-201137

Query Match 90.9%; Score 1387; DB 16; Length 303;
Best Local Similarity 91.1%; Pred. No. 3e-127; Mismatches 16; Indels 2; Gaps 1;
Matches 278; Conservative 9;
QY 1 MEAGGVGLALQARAAGFSGRRHGGIQAPTGSLRVADPAGPAAVAVRARGSKPVAPLRRA 60
DB 1 MEAGGVGLALQTRAAGFSGRRHGGIQAPTGSLRVADPAGPAAVAVRARGSKPVAPLRRA 58
QY 61 KKSOGGHENSHNSVDALLKRSSEVLFYLNRCIYLVGMGSGKSTVGKIMSEVLGYS 120
DB 59 KKSOGGHENSHNSVDALLKRSSEVLFYLNRCIYLVGMGSGKSTVGKIMSEVLGYS 118
QY 121 FFDSDKLVEQAVGMPVSAQIFKVHSEAFPRDNESVLRDLSSMRRLVAVATGGGAVIRPIN 180
DB 119 FFDSDKLVEQAVGMPVSAQIFKVHSEAFPRDNESVLRDLSSMRRLVAVATGGGAVIRPIN 178
QY 181 WRYMKRGLSVMLDVPDLALARRIAKVGTAASRPDLDPQSGDPYTMAFSKLSMLAEQRGDAY 240
DB 179 WKYMKKGLSVMLDVPDLALARRIAKVGTAASRPDLDPQSGDPYTMAFSKLSMLAEQRGDAY 238
QY 241 ANADVRVSLLEETACKQGHDDVSKLTPTDIAIESLHKIESFVIEHTADSSASDAQAESQIQ 300
DB 239 ANADVRVSLLEETACKQGHDDVSKLTPTDIAIESLHKIESFVIEHTADSSASDAQAESQIQ 298
QY 301 RIQTL 305
DB 299 RIQTL 303

RESULT 5
US-10-425-114-62604
Sequence 62604, Application US/10425114
Publication No. US20040034888A1
GENERAL INFORMATION:
APPLICANT: Liu, Jingdong
APPLICANT: Zhou, Yihua
APPLICANT: Kovalic, David K.
APPLICANT: Screen, Steven E.
APPLICANT: Tabaska, Jack E.
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53313)B
CURRENT APPLICATION NUMBER: US/10/425,114
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 73128
SEQ ID NO 62604
LENGTH: 324
TYPE: PRT
ORGANISM: Zea mays subsp. mexicana
FEATURE:
OTHER INFORMATION: Clone ID: UC-ZMFLTEOSINTE77E09_FLI.pap
US-10-425-114-62604

Query Match 90.9%; Score 1387; DB 15; Length 324;
Best Local Similarity 91.1%; Pred. No. 3.3e-127;
Matches 278; Conservative 9; Mismatches 16; Indels 2; Gaps 1;
QY 1 MEAGGVGLALQARAAGFSGRRHGGIQAPTGSLRVADPAGPAAVAVRARGSKPVAPLRRA 60
DB 22 MEAGGVGLALQTRAAGFSGRRHGGIQAPTGSLRVADPAGPAAVAVRARGSKPVAPLRRA 79
QY 61 KKSOGGHENSHNSVDALLKRSSEVLFYLNRCIYLVGMGSGKSTVGKIMSEVLGYS 120
DB 80 KKSOGGHENSHNSVDALLKRSSEVLFYLNRCIYLVGMGSGKSTVGKIMSEVLGYS 139
QY 121 FFDSDKLVEQAVGMPVSAQIFKVHSEAFPRDNESVLRDLSSMRRLVAVATGGGAVIRPIN 180
DB 140 FFDSDKLVEQAVGMPVSAQIFKVHSEAFPRDNESVLRDLSSMRRLVAVATGGGAVIRPIN 199

QY 181 WRYMKRGLSVMLDVPDLALARRIAKVGTAASRPDLDPQSGDPYTMAFSKLSMLAEQRGDAY 240
DB 200 WKYMKKGLSVMLDVPDLALARRIAKVGTAASRPDLDPQSGDPYTMAFSKLSMLAEQRGDAY 259
QY 241 ANADVRVSLLEETACKQGHDDVSKLTPTDIAIESLHKIESFVIEHTADSSASDAQAESQIQ 300
DB 260 ANADVRVSLLEETACKQGHDDVSKLTPTDIAIESLHKIESFVIEHTADSSASDAQAESQIQ 319
QY 301 RIQTL 305
DB 320 RIQTL 324
RESULT 6
US-10-425-114-57834
Sequence 57834, Application US/10425114
Publication No. US20040034888A1
GENERAL INFORMATION:
APPLICANT: Liu, Jingdong
APPLICANT: Zhou, Yihua
APPLICANT: Kovalic, David K.
APPLICANT: Screen, Steven E.
APPLICANT: Tabaska, Jack E.
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53313)B
CURRENT APPLICATION NUMBER: US/10/425,114
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 73128
SEQ ID NO 57834
LENGTH: 326
TYPE: PRT
ORGANISM: Zea mays
FEATURE:
OTHER INFORMATION: Clone ID: UC-ZMFLB73191D05_FLI.pap
US-10-425-114-57834

Query Match 90.9%; Score 1387; DB 15; Length 326;
Best Local Similarity 91.1%; Pred. No. 3.3e-127;
Matches 278; Conservative 9; Mismatches 16; Indels 2; Gaps 1;
QY 1 MEAGGVGLALQARAAGFSGRRHGGIQAPTGSLRVADPAGPAAVAVRARGSKPVAPLRRA 60
DB 24 MEAGGVGLALQTRAAGFSGRRHGGIQAPTGSLRVADPAGPAAVAVRARGSKPVAPLRRA 81
QY 61 KKSOGGHENSHNSVDALLKRSSEVLFYLNRCIYLVGMGSGKSTVGKIMSEVLGYS 120
DB 82 KKSOGGHENSHNSVDALLKRSSEVLFYLNRCIYLVGMGSGKSTVGKIMSEVLGYS 141
QY 121 FFDSDKLVEQAVGMPVSAQIFKVHSEAFPRDNESVLRDLSSMRRLVAVATGGGAVIRPIN 180
DB 142 FFDSDKLVEQAVGMPVSAQIFKVHSEAFPRDNESVLRDLSSMRRLVAVATGGGAVIRPIN 201
QY 181 WRYMKRGLSVMLDVPDLALARRIAKVGTAASRPDLDPQSGDPYTMAFSKLSMLAEQRGDAY 240
DB 202 WKYMKKGLSVMLDVPDLALARRIAKVGTAASRPDLDPQSGDPYTMAFSKLSMLAEQRGDAY 261
QY 241 ANADVRVSLLEETACKQGHDDVSKLTPTDIAIESLHKIESFVIEHTADSSASDAQAESQIQ 300
DB 262 ANADVRVSLLEETACKQGHDDVSKLTPTDIAIESLHKIESFVIEHTADSSASDAQAESQIQ 321
QY 301 RIQTL 305
DB 322 RIQTL 326

RESULT 7
US-10-767-701-44984
Sequence 44984, Application US/10767701
Publication No. US20040172684A1
GENERAL INFORMATION:

```
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(5335)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 44984
; LENGTH: 330
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-C10471_1.pep
US-10-767-701-44984

Query Match      88.6%; Score 1351.5; DB 16; Length 330;
Best Local Similarity 83.3%; Pred. No. 1e-123;
Matches 279; Conservative 8; Mismatches 13; Indels 35; Gaps 3;

QY 1 MEAGGVGLALQAR-----AAGFGSSRRHGGLOAPT 30
   |||||
Db 1 MEA---GLALQRAAGFGSGRRGGQLSPIGSLRVADPAGAAVAAGFGSGRRGGQLSP 57
   |||||

QY 31 GSLRVADPAGPAVARGSKPVAPLRLRAKSSGGHSHNSVDEALLLKRKSEVLFPY 90
   |||||
Db 58 GSLRVADPAGAAVAVARGSKPVAP--LRKKSFGGHENLHNSVDEALLLKRKSEVLFPY 115
   |||||

QY 91 LNGRCIYLVGMGSGKSTVGKIMSEVLGYSPFSDSKLVEQAVCMPSVAQIFKVHSEAPFR 150
   |||||
Db 116 LNGRCIYLVGMGSGKSTVGKIMSEVLGYSPFSDSKLVEQAVCMPSVAQIFKVHSEAPFR 175
   |||||

QY 151 DNESSVRLDSSMRRLVATGGAVIRPINWRYMKRGLSVWLDVPLDALARRIAKVGTAS 210
   |||||
Db 176 DNESSVRLDSSMRRLVATGGAVIRPINWRYMKRGLSVWLDVPLDALARRIAKVGTAS 235
   |||||

QY 211 RPLDQPSGDPYAMAFSKLSMLAQORDAYANADVRSLEETACKQGHDDVSKLTPTDIA 270
   |||||
Db 236 RPLDQPSGDPYTMAFSKLSMLAEQORDAYANADVRSLEETACKQGHDDVSKLTPTDIA 295
   |||||

QY 271 IESLHKIESFVIEHTADSSASDAQAESQIQRIOTL 305
   |||||
Db 296 IESLHKIESFVSEHTPDNPSDSQAESQIQRIOTL 330
   |||||

RESULT 8
US-10-437-963-147803
; Sequence 147803, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 147803
; LENGTH: 308
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_48297C.1.pep
US-10-437-963-147803
```

```
Query Match      85.3%; Score 1301.5; DB 16; Length 308;
Best Local Similarity 86.2%; Pred. No. 7.4e-119;
Matches 268; Conservative 13; Mismatches 21; Indels 9; Gaps 4;

QY 1 MEAGGVGLALQARAGF--GSSRRHGGGL-----QAPTGSLRVADPAGPAVAVRARGSKPVA 54
   |||||
Db 1 MEA-CVGLALQSRAGFGSDRRRSALYGGEGRARIGSLRVAEPAAKAAVWARGSKPVA 59
   |||||

QY 55 PLRLRAKSSGGHSHNSVDEALLLKRKSEVLFPYLNRCIYLVGMGSGKSTVGKIMS 114
   |||||
Db 60 P--LRAKKSSGGHETLHNSVDEALLLKRKSEVLFPYLNRCIYLVGMGSGKSTVGKIMS 117
   |||||

QY 115 EVLGYSPFSDSKLVEQAVCMPSVAQIFKVHSEAPFRDNESSVRLDSSMRRLVATGGGA 174
   |||||
Db 118 EVLGYSPFSDSKLVEQAVCMPSVAQIFKVHSEAPFRDNESSVRLDSSMRRLVATGGGA 177
   |||||

QY 175 VIRPINWRYMKRGLSVWLDVPLDALARRIAKVGTASRPLDQPSGDPYAMAFSKLSMLAQ 234
   |||||
Db 178 VIRPINWRYMKRGLSVWLDVPLDALARRIAKVGTASRPLDQPSGDPYTMAFSKLSMLAE 237
   |||||

QY 235 QRGDAYANADVRSLEETACKQGHDDVSKLTPTDIAIESLHKIESFVIEHTADSSASDAQ 294
   |||||
Db 238 QRGDAYANADVRSLEETACKQGHDDVSKLTPTDIAIESFHKIENFVIEHTVDNVPVGSQ 297
   |||||

QY 295 AESQIQRIOTL 305
   |||||
Db 298 ADSRAQRIOTL 308
   |||||

RESULT 9
US-10-660-226-28
; Sequence 28, Application US/10660226
; Publication No. US20040064848A1
; GENERAL INFORMATION:
; APPLICANT: Cahoon, Rebecca E.
; APPLICANT: Falco, S. Carl
; APPLICANT: Famodu, Layo O.
; APPLICANT: Hitz, William D.
; APPLICANT: Rendina, Alan
; TITLE OF INVENTION: Chorismate Biosynthesis Enzymes
; FILE REFERENCE: BB-1159-C
; CURRENT APPLICATION NUMBER: US/10/660,226
; CURRENT FILING DATE: 2003-09-11
; PRIOR APPLICATION NUMBER: US/09/354,501
; PRIOR FILING DATE: 1999-07-16
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/093,611
; PRIOR FILING DATE: EARLIER FILING DATE: July 21, 1998
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 28
; LENGTH: 273
; TYPE: PRT
; ORGANISM: Triticum aestivum
US-10-660-226-28

Query Match      76.0%; Score 1159; DB 15; Length 273;
Best Local Similarity 85.5%; Pred. No. 6e-105;
Matches 235; Conservative 17; Mismatches 17; Indels 6; Gaps 3;

QY 32 SLRVADPAGPAVAVRARGSKPVAPLRLRAKSS--GGHSHNSVDEALLLKRKSEVLFPY 90
   |||||
Db 4 SLRVSDLVGSPAAVARGAKPVVP--LRAKSSGGGHENLHNSVDDALLLKRKSEVLFPQ 61
   |||||

QY 91 LNGRCIYLVGMGSGKSTVGKIMSIVLGYSPFSDSKLVEQAVCMPSVAQIFKVHSEAPFR 150
   |||||
Db 62 LNGRCIYLVGMGSGKSTVGKILAEVLGYSPFSDSKLVEQAVCMPSVAQIFKVHSEAPFR 121
   |||||

QY 151 DNESSVRLDSSMRRLVATGGAVIRPINWRYMKRGLSVWLDVPLDALARRIAKVGTAS 210
   |||||
Db 122 DNESSVRLDSSMRRLVATGGAVIRPINWRYMKRGLSVWLDVPLEALARRIAKVGTAS 181
   |||||

QY 211 RPLDQPSGDPYAMAFSKLSMLAQORDAYANADVRSLEETACKQGHDDVSKLTPTDIA 270
   |||||
```

```
Db 182 RPLDQPSGDPYTMAFSKLSMLAEQRGDAYANADVRSLEETASKLGHDDVSKLTPIDIA 241
Qy 271 IESLHKIESFVIEHTADSSASDAQESQIQRIQTL 305
Db 242 LESLHKIESFVVE---DTAVADSQIESQSQRMHTL 273

RESULT 10
US-10-425-115-201136
; Sequence 201136, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 201136
; LENGTH: 217
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(217)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_115021C.1.pap
US-10-425-115-201136

Query Match 69.5%; Score 1061; DB 16; Length 217;
Best Local Similarity 98.6%; Pred. No. 1.8e-95;
Matches 212; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 MEAGGVGLALQARAAFGFSRRHGGIQAPTGSLRVADPAGPAVAVRARGSKPAPLRRA 60
Db 1 MEAGGVGLALQARAAFGFSRRHGGIQAPTGSLRVADPAGPAVAVRARGSKPAPLRRA 60
Qy 61 KXSSGSHNSHNSVDEALLKKRSEVLFLYNGRCIYLVGMMSGKSTVGKIMSEVLGYS 120
Db 61 KXSSGSHNSHNSVDEALLKKRSEVLFLYNGRCIYLVGMMSGKSTVGKIMSEVLGYS 120
Qy 121 FPDSDKLVEQAVGMPMSVAQIFKVHSEAFPRDNHSSVLRDLSSMRRLVWATGGAVIRPIN 180
Db 121 FPDSDKLVEQAVGMPMSVAQIFKVHSEAFPRDNHSSVLRDLSSMRRLVWATGGAVIRPIN 180
Qy 181 WRYMKRGLSVWLDVPLDALARRIAKVGTASRPLLD 215
Db 181 WRYMKRGLSVWLDVPLDALARRIAKVGTASRPLLD 215

RESULT 11
US-10-437-963-187425
; Sequence 187425, Application US/10437963
; Publication No. US2004023343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
```

```
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 187425
; LENGTH: 311
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_84129C.1.pap
US-10-437-963-187425

Query Match 58.1%; Score 886; DB 16; Length 311;
Best Local Similarity 59.8%; Pred. No. 4.5e-78;
Matches 202; Conservative 27; Mismatches 49; Indels 60; Gaps 9;

Qy 1 MEAGGVGLALQARAA-GFGSS---RHRGGLQAPTG-SLRVADPAG-----PAVA 44
Db 1 MEA-RAGLAMQRAAAGVGAGPGVGRGRARDPRGEAANRGPAGRRPRRTGGGQAPRA 59
Qy 45 VRARG-----SKVPAPLRRAKSS-----GGHSHNSVDEALLKKRSEVL 88
Db 60 VLPQGVQSFLIGLGEVVTFFVWISKDNLGVNGVVGQHDLSHNSVDEALLKKRSEVL 119
Qy 89 FYLNGRCIYLVGMMSGKSTVGKIMSEVLGYSFFDSKLVQAVGMPMSVAQIFKVHSEAF 148
Db 120 FYLNGRCIYLVGMMSGKSTVAKILAEVLGYSFFDSKLVQAVGMPMSVAQIFKVHSEAF 179
Qy 149 FRDNHSSVLRDLSSMRRLVWATGGAVIRPINWRYMKRGLSVWLDVPLDALARRIAKVGT 208
Db 180 FRDNEA-----MKYMKKGLSVWLDVPLDALARRIAQVGT 213
Qy 209 ASRPLDQPSGDPYAMAFSKLSMLAQRGDAYANADVRSLEETACKQGHDDVSKLTPTD 268
Db 214 ASRPLDQPSGDPYTAAFSKLSMLAEQRGDAYANADARVSLSEETAAKQGHDDVSKLTPTD 273
Qy 269 IAIESLHKIESFVIEH-TADSSASDAQESQIQRIQTL 305
Db 274 IAIEALLKIENFVTEHSTSSGPGVDLVDQSQRRTKAL 311

RESULT 12
US-10-660-226-26
; Sequence 26, Application US/10660226
; Publication No. US20040064848A1
; GENERAL INFORMATION:
; APPLICANT: Cahoon, Rebecca E.
; APPLICANT: Falco, S. Carl
; APPLICANT: Famodu, Lavo O.
; APPLICANT: Hitz, William D.
; APPLICANT: Rendina, Alan
; TITLE OF INVENTION: Chorismate Biosynthesis Enzymes
; FILE REFERENCE: BB-1159-C
; CURRENT APPLICATION NUMBER: US/10/660,226
; CURRENT FILING DATE: 2003-09-11
; PRIOR FILING DATE: 1999-07-16
; PRIOR APPLICATION NUMBER: US/09/354,501
; PRIOR FILING DATE: EARLIER FILING DATE: July 21, 1998
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 26
; LENGTH: 282
; TYPE: PRT
; ORGANISM: Triticum aestivum
US-10-660-226-26

Query Match 57.3%; Score 874; DB 15; Length 282;
Best Local Similarity 65.2%; Pred. No. 5.9e-77;
Matches 180; Conservative 33; Mismatches 49; Indels 14; Gaps 2;

Qy 22 HRGGLQAPTGSLRVADPAG-----GPVAVRARGSKPAPLRRAKSS-----SGGH 67
Db 1 HEAKRKQGGFPFATVPVAVRLDQNPAPRLVLRDTAGSRSTDPIRGASLKALCCHKSAGT 60
Qy 68 ENSHNSVDEALLKKRSEVLFLYNGRCIYLVGMMSGKSTVGKIMSEVLGYSFFDSKLTPTD 127
```

Db 61 EXAHTSADAEALVKQAEVDLPVNDRCVYLVGMGSGKTTVGKIIAEVLGYSFDFSDKL 120
QY 128 VEQAVGMPVAQIFKVHSEAFPRDNESVLRDLSSMRRLVATGGGAVIRPINWYMRKG 187
Db 121 VEQSGVGPVABIFQVHSEAFPRDNESVLRDLSSMRRLVATGGGAVIRPINWYMRKG 180
QY 188 LSVMLDVPDLALARRIAVGTASRPLLDQPSGDPYAMAFSKLSMLAQOQGDYANADVRV 247
Db 181 LTIWLDVPLDALARRIAVGTASRPLLDQPSGDPYAMAFSKLSMLAQOQGDYANADVRV 240
QY 248 SLEEIAKQGHDDVSKLTPTDIAIESLHKIESFVIE 283
Db 241 SLENTAFKQGHDDVNVLTSAIAIEALLKQESFLTE 276

RESULT 13
US-10-660-226-20
; Sequence 20, Application US/10660226
; Publication No. US20040064848A1
; GENERAL INFORMATION:
; APPLICANT: Cahoon, Rebecca E.
; APPLICANT: Falco, S. Carl
; APPLICANT: Famodu, Layo O.
; APPLICANT: Hitz, William D.
; APPLICANT: Rendina, Alan
; TITLE OF INVENTION: Chorismate Biosynthesis Enzymes
; FILE REFERENCE: BB-1159-C
; CURRENT APPLICATION NUMBER: US/10/660,226
; PRIOR FILING DATE: 2003-09-11
; PRIOR APPLICATION NUMBER: US/09/354,501
; PRIOR FILING DATE: 1999-07-16
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/093,611
; PRIOR FILING DATE: EARLIER FILING DATE: July 21, 1998
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 20
; LENGTH: 245
; TYPE: PRT
; ORGANISM: Sorghum
US-10-660-226-20

Query Match 56.4%; Score 860; DB 15; Length 245;
Best Local Similarity 70.4%; Pred. No. 1.1e-75;
Matches 174; Conservative 32; Mismatches 33; Indels 8; Gaps 2;
QY 39 AGPAVAVARGSKPVAPLRRAKSSGGHSHNSVDEALLKRSSEVLFYINGRCIYL 98
Db 3 AGPAL-----RP-AKLAVSCSAKSAGTGKVHYSTDEALILQKQAQDVLPLDGRCVYL 54
QY 99 VGMGSGKSTVGKIMSEVLGYSFDFSDKLVEQAVGMPVAQIFKVHSEAFPRDNESVLR 158
Db 55 VGMGSGKSTVGKIIAEVLGYSFDFSDKLVEQAVGMPVAQIFKVHSEAFPRDNESVLR 114
QY 159 DLSSMRRLVATGGGAVIRPINWYMRKLSVWLDVPLDALARRIAVGTASRPLLDQPS 218
Db 115 DLSSMRRLVATGGGAVIRPINWYMRKLSVWLDVPLDALARRIAVGTASRPLLDQPS 174
QY 219 GOPYAMAFSKLSMLAQOQGDYANADVRVLSLEIAKQGHDDVSKLTPTDIAIESLHKIE 278
Db 175 GOPYAMAFSKLSVWLDVPLDALARRIAVGTASRPLLDQPSGDPYAMAFSKLSMLAQOQ 234
QY 279 SFVIEHT 285
Db 235 SFLTEXT 241

RESULT 14
US-10-425-115-253534
; Sequence 253534, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.

; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 253534
; LENGTH: 292
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_162804C.1.pep
US-10-425-115-253534

Query Match 56.2%; Score 858; DB 16; Length 292;
Best Local Similarity 66.8%; Pred. No. 2.3e-75;
Matches 173; Conservative 34; Mismatches 52; Indels 0; Gaps 0;
QY 27 QAPTGSRLVADPAGPAVAVARGSKPVAPLRRAKSSGGHSHNSVDEALLKRSSE 86
Db 30 RVPTARLAVAAADRPRLVLLGADTERRAADPALRCAAQASAGTGKVHYSDADALILQKQAQD 89
QY 87 VLFYINGRCIYLVGMGSGKSTVGKIMSEVLGYSFDFSDKLVEQAVGMPVAQIFKVHSE 146
Db 90 VLPYLDGRVCYLVGMGSGKTTVGKILSEVLGYSFDFSDKLVEQAVGMPVAQIFKVHSE 149
QY 147 AFFRDNESVLRDLSSMRRLVATGGGAVIRPINWYMRKLSVWLDVPLDALARRIAK 206
Db 150 TFFRDNESVLRDLSSMRRLVATGGGAVIRPINWYMRKLSVWLDVPLDALARRIAK 209
QY 207 GTASRPLLDQPSGDPYAMAFSKLSMLAQOQGDYANADVRVLSLEIAKQGHDDVSKLT 266
Db 210 GTASRPLLDQPSGDPYAMAFSKLSMLAQOQGDYANADVRVLSLEIAKQGHDDVSKLT 269
QY 267 TDIAIESLHKIESFVIEHT 285
Db 270 STIAIEALLKQESFLTEXT 288

RESULT 15
US-10-660-226-12
; Sequence 12, Application US/10660226
; Publication No. US20040064848A1
; GENERAL INFORMATION:
; APPLICANT: Cahoon, Rebecca E.
; APPLICANT: Falco, S. Carl
; APPLICANT: Famodu, Layo O.
; APPLICANT: Hitz, William D.
; APPLICANT: Rendina, Alan
; TITLE OF INVENTION: Chorismate Biosynthesis Enzymes
; FILE REFERENCE: BB-1159-C
; CURRENT APPLICATION NUMBER: US/10/660,226
; CURRENT FILING DATE: 2003-09-11
; PRIOR APPLICATION NUMBER: US/09/354,501
; PRIOR FILING DATE: 1999-07-16
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/093,611
; PRIOR FILING DATE: EARLIER FILING DATE: July 21, 1998
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 12
; LENGTH: 231
; TYPE: PRT
; ORGANISM: Zea mays
US-10-660-226-12

Query Match 55.1%; Score 841; DB 15; Length 231;
Best Local Similarity 73.7%; Pred. No. 7.6e-74;
Matches 165; Conservative 31; Mismatches 28; Indels 0; Gaps 0;
QY 62 KSSGGHSHNSVDEALLKRSSEVLFYINGRCIYLVGMGSGKSTVGKIMSEVLGYSP 121

Db	4	Q	S	A	G	G	T	K	V	H	S	A	D	D	A	L	I	L	Q	K	A	O	D	V	L	P	Y	L	D	G	R	C	V	Y	L	V	G	M	G	S	G	K	T	T	V	G	K	I	L	S	E	V	L	G	Y	S	F	63				
Qy	122	F	D	S	D	K	L	V	E	O	A	V	G	M	P	S	V	A	Q	I	F	K	V	H	S	E	A	F	F	R	D	N	E	S	S	V	L	R	D	L	S	S	M	R	R	L	V	V	A	T	G	G	G	A	V	I	R	P	I	N	W	181
Db	64	F	D	S	D	K	L	V	E	K	A	V	G	I	S	S	V	A	E	I	F	Q	L	H	S	E	T	F	F	R	D	N	E	S	E	V	L	T	D	L	S	S	M	H	R	L	V	V	A	T	G	G	G	A	V	I	R	P	I	N	W	123
Qy	182	R	Y	M	K	R	G	L	S	V	W	L	D	V	P	L	D	A	L	A	R	R	I	A	K	V	G	T	A	S	R	P	L	L	D	Q	P	S	G	D	P	Y	A	M	A	F	S	K	L	S	M	L	A	Q	O	R	G	D	A	Y	A	241
Db	124	S	Y	M	K	G	L	T	V	W	L	D	V	P	L	D	A	L	A	R	R	I	A	A	V	G	T	A	S	R	P	L	L	H	Q	E	S	G	D	P	Y	A	K	A	Y	A	K	L	T	S	L	F	E	Q	R	M	D	S	Y	A	183	
Qy	242	N	A	D	V	R	S	L	E	E	I	A	C	K	O	G	H	D	V	S	K	L	T	P	T	D	I	A	E	S	L	H	K	I	E	S	F	V	I	E	H	T	285																			
Db	184	N	A	D	A	R	V	S	L	E	H	I	A	L	K	O	G	H	N	D	V	T	I	L	T	P	S	T	I	A	I	E	A	L	L	K	M	E	S	F	L	T	E	K	T	227																

Search completed: August 25, 2005, 11:19:05
Job time : 67 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 25, 2005, 19:08:37 ; Search time 4914 Seconds
(without alignments)

3007.497 Million cell updates/sec

Title: US-10-660-226-10

Perfect score: 1526

Sequence: 1 MEAGVGGLQARAAGFGSS.....ADSSASDAQESQIQRIQTL 305

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5

Ygapop 10.0 , Ygapext 0.5

Fgapop 6.0 , Fgapext 7.0

Delop 6.0 , Delext 7.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n model -DEV=xlh

-Q=/cgn2_1/USFTO_spool/US10660226/runat_25082005_110427_28529/app_query.fasta_1.455

-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPECL=0 -LOOPEXT=0

-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45

-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL

-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000

-USER=US10660226 @cgn 1 1 3731 @runat_25082005_110427_28529 -NCFU=6 -ICPU=3

-NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG

-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6

-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*

1: gb.ba.*

2: gb.htg.*

3: gb.in.*

4: gb.om.*

5: gb.ov.*

6: gb.pat.*

7: gb.ph.*

8: gb.pl.*

9: gb.pr.*

10: gb.ro.*

11: gb.ste.*

12: gb.sv.*

13: gb.un.*

14: gb.vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1162.5	76.2	1356	8 AK066687	Oryza sat
2	1159	76.0	1061	8 BT009529	Triticum
3	1025.5	67.2	1918	8 AK109730	Oryza sat
4	881.5	57.8	1323	8 BT009514	Triticum

C	5	836.5	54.8	111777	2	AP004089	Oryza sat
	6	836.5	54.8	154796	8	AP004774	Oryza sat
	7	836.5	54.8	155168	8	AP005848	Oryza sat
C	8	733	48.0	137319	8	AP003513	Oryza sat
	9	697.5	45.7	2127	8	AK067948	Oryza sat
	10	691	45.3	1167	8	AK070318	Oryza sat
	11	655	42.9	1300	8	AK118899	Arabidops
	12	642.5	42.1	1293	8	LESHKINP	X63560 L.esculentu
	13	642.5	42.1	1303	8	BT012939	Lycopersi
	14	606.5	39.7	903	8	BT005291	Arabidops
	15	606.5	39.7	1294	8	AK117791	Arabidops
C	16	516.5	33.8	126659	8	OSJUN00015	Arabidops
C	17	516.5	33.8	155576	2	OSJUN00094	Oryza sat
	18	459.5	30.1	107848	8	AC007019	Arabidops
	19	440	28.8	1016	8	AY596190	Cucumis s
	20	390	25.6	129047	8	ATF23K16	Arabidops
	21	390	25.6	198151	8	ATCHRIV91	Arabidops
	22	374	24.5	309	6	AX315616	Sequence
	23	357	23.4	297850	1	AP006577	Gloebact
	24	340.5	22.3	306803	1	AE017161	Prochloro
	25	340.5	22.3	342050	1	AP003585	Nostoc sp
	26	331.5	21.7	843	8	BT012407	Arabidops
	27	331.5	21.7	898	8	BT012132	Arabidops
	28	321.5	21.1	344615	1	BX569695	Synechoco
C	29	321	21.0	137740	1	D90900	Synechocyst
	30	311.5	20.4	300450	1	AP005371	Thermosyn
	31	301	19.7	110000	2	AP006502_02	Continuation (3 of
C	32	285	18.7	349391	1	BX572095	Prochloro
	33	272.5	17.9	349742	1	BX572090	Prochloro
C	34	268.5	17.6	300600	1	AP005935	Bradyrhiz
C	35	262	17.2	110000	1	AE017180_22	Continuation (23 o
C	36	256.5	16.8	348971	1	BX572594	Rhodospseu
	37	254.5	16.7	110000	1	BX897699_18	Continuation (19 o
	38	254	16.6	5062	14	SVU27645	Sealth vir
C	39	252.5	16.5	110000	1	AP006840_20	Continuation (21 o
	40	250	16.4	81384	1	AE007700_15	Continuation (16 o
	41	249.5	16.3	15580	1	AE000776	Aquifex a
	42	246	16.1	1125	8	AK064812	Oryza sat
	43	245	16.1	696	6	AR449503	Sequence
	44	245	16.1	14335	6	AR408732	Sequence
	45	245	16.1	14335	6	AX067436	Sequence

ALIGNMENTS

RESULT 1	AK066687	Oryza sativa (japonica cultivar-group)	1356 bp	mRNA	linear	PLN 24-JUL-2003
LOCUS	AK066687	Oryza sativa (japonica cultivar-group)	1356 bp	mRNA	linear	PLN 24-JUL-2003
DEFINITION	AK066687	Oryza sativa (japonica cultivar-group)	1356 bp	mRNA	linear	PLN 24-JUL-2003
ACCESSION	AK066687.1	GI:32976705				
VERSION	AK066687.1	GI:32976705				
KEYWORDS	FLI_CDNA; CAP trapper.					
SOURCE	Oryza sativa (japonica cultivar-group)					
ORGANISM	Oryza sativa (japonica cultivar-group)					
REFERENCE	1	Oryza sativa (japonica cultivar-group)				
AUTHORS	1	Oryza sativa (japonica cultivar-group)				
	1	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.				

The Rice Full-Length cDNA Consortium, National Institute of Agricultural Sciences Rice Full-Length cDNA Project Team: Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T., Foundation of Advancement of International Science Genome Sequencing & Analysis Group: Ohtsuki, Y., Murakami, K., Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda, Y., Kurosaki, T., Kodama, T., Masuda, H., Kobayashi, M., Xie, Q., Lu, M., Narikawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Niihara, J., Ikeda, R., Ishibiki, J., Kawamata, M., Yoshimura, A., Miura, J., Kusumegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., KIKEN: Kawai, J., Carninci, P., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashidume, W., Hayatsu, N., Imotani, K., Ishii, Y., Itoh, M.,

Kagawa, I., Kondo, S., Konno, H., Miyazaki, A., Osato, N., Ota, Y.,
Saito, R., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T.,
Yoshino, M. and Hayashizaki, Y.

TITLE Collection, mapping, and annotation of over 28,000 cDNA clones from japonica rice

JOURNAL
Science 301 (5631), 376-379 (2003)

MEDLINE 22752273

12869764
PUBMED

FORMED
REFERENCE 2 (pages 1 to 1356)

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K.,
Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W.,
Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hiraoka, T.,
Hori, F., Hotta, I., Iida, J., Iida, Y., Ikeda, R., Imamura, K.,
Imotani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I.,
Kanagawa, S., Katoh, H., Kawagaeshira, N., Kawai, J., Kawamata, M.,
Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M.,
Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Kurotaki, T., Kusumegi, T., Li, C., Lu, M.,
Maeda, H., Matsubara, K., Matsuyama, T., Miura, J., Miyazaki, A.,
Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakamura, M.,
Naniki, T., Narikawa, R., Niikura, J., Nishi, K., Nomura, K.,
Numasaki, R., Ohneda, E., Ohno, M., Ohtsuki, K., Oka, M., Ooka, H.,
Osato, N., Ota, Y., Otomo, Y., Ryu, R., Saitoh, H., Sakai, C., Sakai, K.,
Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K., Shibata, K.,
Shinagawa, A., Shiraki, T., Shishik, T., Sogabe, Y., Sugano, S.,
Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami-Takeda, Y.,
Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A.,
Toya, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yahagi, W.,
Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S. and
Yoshimura, A.

Direct Submission

Submitted (03-DEC-2001) Shoshi Kikuchi, National Institute of
Agrobiological Sciences, Department of Molecular Genetics, Head of
Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki
305-8660, Japan (E-mail: skiki@nias.affrc.go.jp,
Tel: 81-29-838-7007 Fax: 81-29-838-7007)

COMMENT
This clone is one of the 28K full-length cDNA clones from japonica rice.

URL: <http://cdna01.dna.affrc.go.jp/cDNA/>

URL: <http://comair.cmu.edu/90.jp/CowM/>
 NIRS Rice Full-Length cDNA Project Team: Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T. and Yamamoto, M.

FAIS Genome Sequencing & Analysis Group: Otomo, Y., Iida, Y., Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M., Kodama, T., Kurosaki, T., Kusumegi, T., Lu, M., Masuda, H., Miura, J., Mizuno, K., Nariakawa, R., Niikura, J., Oka, M., Ryu, R., Sugano, S., Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S., Yoshimura, A., Matsubara, K. and Murakami, K.

Genome Exploration Research Group in Riken Genomic Sciences Center
and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K.,
Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T.,
Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K.,
Hiraoka, T., Hori, F., Iida, J., Imamura, K., Imotani, K., Ishii, Y.,
Itoh, M., Kawaga, I., Kanagawa, S., Katoh, H., Kawai, J.,
Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Numaaki, R., Ohno, M., Oshio, N.,
Ota, Y., Saitho, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H.,
Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F.,
Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K.,
Yasukuni-Akai, A. and Hayashizaki, Y.

Location/Qualifiers

1. .1356

```

2: 1530
/organism="Oryza sativa (japonica cultivar-group)"

```

```
/organism="Oryza  
/mol_type="rRNA"
```

```
/mol_type="RNA"
/cu1 fixar="Nipponbare"
```

```
/culcivai=Nippomare  
/db_yref="taxon:39947"
```

```
/dev_xlivel=lx011:3999
/c1one="J013074E04"
```

ORIGIN

Alignment Scores:

Pred. No.:	1.55e-78	Length:	1356
Score:	1162.50	Matches:	259
Percent Similarity:	71.02%	Conservative:	13
Best Local Similarity:	67.62%	Mismatches:	22
Query Match:	76.18%	Indels:	90
DB:	8	Gaps:	6

US-10-660-226-10 (1-305) x AK066687 (1-1356)

Qy	1	MetGluAlaGlyGlyValGlyLeuAlaLeuGlnAlaAArgAlaGlyPhe----	GlySer	19
Db	81	ATGAGGCG---GGCGTGGGGCTGGCGCTCGACTCGCGGGCGGGGGTTCGGCGGGCTCC		137
Qy	20	SerArgHisArgGlyGlyLeu-----GlnAlaProThrGlySerLeuArg		34
Db	138	GACCGCGCGCGAGCGCTCTACGGCGGCGGGCGCGCGAGATCGGGAGCTTGAGG		197
Qy	35	ValAlaAspProAlaGlyProAlaValAlaValAArgGlySerLysProValAla		54
Db	198	GTGCTGAGCGCGGGTGGCGAAGCGCGTGTGTGGCTCGCGGGTCCAAGCCGGTCCGC		257
Qy	55	ProLeuArgLeuArgAlaLysLysSerSerGlyHisGluAsnSerHisAsnSerVal		74
Db	258	CCG-----CTCCGTGCAAGAAATCGTCGGAGGTCATGAAACATTGCAATACTCGGTT		311
Qy	75	AspGluAlaLeuLeuLysArgLysSerGluGluValLeuPheTyrLeuAsnGlyArg		94
Db	312	GATGAAGCCCTCTTGCTAAGAGAGAAATCAGAGAAGTCTCTTCTATTGTAATGACCGG		371
Qy	95	CysIleTyrLeuValGlyMetMetGlySerGlyLysSerThrValGlyLysIleMetSer		114
Db	372	TGTATTACCTAGTTGGATGATGGGTTCTGGAAAGAAGTACTGTGGGAAGATCATGTCT		431
Qy	115	GluValLeuGlyTyrSerPhePheAspSerAspLysLeuValGluGlnAlaValGlyMet		134
Db	432	GAAGTTTGGGTTATTTCGTTCTTTGTATGTATGATAAAATTTGTCGAACAAGCTGTGGGCATG		491
Qy	135	ProSerValAlaGlnIlePheLysValHisSerGluAlaPhePheArgAspAsnGluSer		154
Db	492	CCTTCAGTCGCTCAAAATTTTCAAGGTTCAATAGTGAAGCCCTTCTTTAGGGGATAATGAGAGT		551
Qy	155	SerValLeuArgAspLeuSerSerMetArgArgLeuValValAlaThrGlyGlyAla		174
Db	552	AGTGTCTTGAGGATTTGCTCTCAATGAAGCGATTAGTTGTTGCTACTCGAGGTGGTGTCT		611
Qy	175	ValIleArgProIleAsnTrpArgTyrMetLysArgGlyLeuSerValTrpLeuAspVal		194
Db	612	GTTATCCGACGACTTAACGTGAATAATCATGAAGAAGGC-----		650
Qy	195	ProLeu-AspAlaLeuAlaArgArgIleAlaLysValGlyThrAlaSerArgProLeuLe		214
Db	651	---CTTGGACGCTCTTGCTAGGCGTATGCTAAAGTGGGACATGCCTCCCGTCCCTCTCT		707
Qy	214	uAspGlnProSerGlyAspProTyrAlaMetAla-----		225
Db	708	AGATCAACCATCTGGTGATCCATACAAATGGT-GACTACTTGGCCCATGAATTTATGCC		766
Qy	225	-----		225
Db	767	ATGCATGAATATGATATTATTTTCAGATGTATTTTTTTAAATTTTTTAATCAGTCAATATAG		826
Qy	225	-----		225
Db	827	GCTCATTCATTCAATGTPAGAAATGAAAACTGCAAGTATGATCCTGCACATCTGCACAAC		886
Qy	225	-----		225
Db	887	TGGTAGTGAATTTCTTTTCATTTTCAGTCTAGTATGAGTTGCTCACTGGATTCTTCCTATC		946
Qy	226	-----PheSerLysLeuSerMetLeuAlaGlnGlnArgGlyAspAlaTyrAlaAsnAl		243
Db	947	TTAGGCTTTTCTTAAACTCAGCATGCTCGCGGAGCAAGGGCGATGCTTATGCAATGC		1006


```

QY 243 aAspValArgValSerLeuGluIleAlaCysLysGlnGlyHisAspValSerLy 263
Db 1007 TGATGTGAGGGTTCTCTTGAAGAGATTGCATCTAAACAGGGTCATGATGATCGAA 1066
QY 263 sLeuThrProThrAspIleAlaIleGluSerLeuHisLysLysPheValIleG 283
Db 1067 GCTAACACCAACATGACATTCGTTAGTTCGTTTCATAGATGAGAACTTTGTCA 1126
QY 283 uHisThrAlaAspSerSerAlaSerAspAlaGlnAlaGluSerGlnIleArgIleG 303
Db 1127 ACATACCGTTGACAAATCCGGTGGTGGTCCACAGGCTGACTCAGTCTCAGAG 1186
QY 303 nThrLeu 305
Db 1187 GACCTTG 1193

RESULT 2
LOCUS BT009529 1061 bp mRNA linear PLN 20-JUN-2003
DEFINITION Triticum aestivum clone wrl.pk0122.a3:fis, full insert mRNA
sequence.
ACCESSION BT009529
VERSION BT009529.1 GI:32129080
KEYWORDS FLI_CDNA.
SOURCE Triticum aestivum (bread wheat)
ORGANISM Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poideae; Triticeae; Triticum.
REFERENCE 1 (bases 1 to 1061)
AUTHORS Tingey,S.V., Walters,P., Powell,W., Dolan,M., Miao,G.-H.,
Caraher,N.R., Hanafey,M.K. and Hainey,C.F.
TITLE Direct Submission
JOURNAL Submitted (20-JUN-2003) Crop Genetics, E. I. DuPont de Nemours and
Company, 1 Innovation Way, P.O. Box 6104, Newark, DE 19714-6104,
USA

FEATURES
Source
1..1061
Location/Qualifiers
1..1061
/organism="Triticum aestivum"
/mol_type="mRNA"
/db_xref="taxon:4565"
/clone="wrl.pk0122.a3:fis"

ORIGIN
Alignment Scores:
Pred. No.: 2,14e-78 Length: 1061
Score: 1159.00 Matches: 235
Percent Similarity: 91.64% Conservative: 17
Best Local Similarity: 85.45% Mismatches: 17
Query Match: 75.95% Indels: 6
DB: 8 Gaps: 3

US-10-660-226-10 (1-305) x BT009529 (1-1061)
QY 32 SerLeuArgValAlaAspProAlaGlyProAlaValAlaValArgAlaArgGlySerLys 51
Db 11 AGCTTCGGTGTAGTATCTGGTGGGTCGCCGCCGCGTGGCGCGCGCGCGCGCGCAAG 70
QY 52 ProValAlaProLeuArgLeuArgAlaLysLysSerSer---GlyGlyHisGluAsnSer 70
Db 71 CCGGTGCTCCG-----CTCCGCGCAAGAAATCGTCTGGAGGAGTCTATGAGAAGTTG 124
QY 71 HisAsnSerValAspGluAlaLeuLeuLysArgLysSerSerGluValLeuPheTyr 90
Db 125 CATAACTCCGTTGACATGCTCTTGTGGTGGAGAGAAATCAGAGAGGTCTTTTCCAG 184
QY 91 LeuAsnGlyArgCysIleTyrLeuValGlyMetMetGlySerGlyLysSerThrValGly 110
Db 185 TTGAACGGTCGGTGCATCTACCTAGTTGGAATGATGGTTCGGGGNAAAGTACGGTGGGG 244
QY 111 LysIleMetSerGluValLeuGlyTyrSerPheAspSerAspLysLeuValGluGln 130

```

```

Db 245 AAGATCTTGGCTGAAGTTTGGGTATTTCATCTCTTCGACAGTGATAAATTTGGTCGAACAA 304
QY 131 AlaValGlyMetProSerValAlaGlnIlePheLysValHisSerGluAlaPhePheArg 150
Db 305 GCTGTTGGCATGCTTCAGTTGCTCAAAATTTTCAAGGTTTCATAGTGAAGCCTTCTTCAGA 364
QY 151 AspAsnGluSerSerValLeuArgAspLeuSerSerMetArgArgLeuValValAlaThr 170
Db 365 GATAATGAGAGTAGTGTCTTGAGGGAATTTGTCTCAATCGCGGATTTAGTTTGTGCTACT 424
QY 171 GlyGlyGlyAlaValIleArgProIleAsnTyrArgTyrMetLysArgGlyLeuSerVal 190
Db 425 GGAGTGGTGGCTGTTATCCGACCGTTAACTGAAAAAATATGAAGAGGGCCTATCTGT 484
QY 191 TrpLeuAspValProLeuAspAlaLeuAlaArgArgIleAlaLysValGlyThrAlaSer 210
Db 485 TGGTTGGATGTGCCCTTGGAGCTCTTGCNAGGCGTATTGCTAAAGTGGGACTGCCTCG 544
QY 211 ArgProLeuLeuAspGlnProSerGlyAspProTyrAlaMetAlaPheSerLysLeuSer 230
Db 545 CGTCTCTTCTAGATCAACCATCCGGTGATCCATACACAATGGCCTTTTCGAAACTCAGC 604
QY 231 MetLeuAlaGlnArgGlyAspAlaTyrAlaAsnAlaAspValArgValSerLeuGlu 250
Db 605 ATGCTCGCGGAGCAAAAGGGGGGATGCTTATGCAAAATGCTGATGTGAGAGTTTCTTCGAA 664
QY 251 GluIleAlaCysLysGlnGlyHisAspValSerLysLeuThrProThrAspIleAla 270
Db 665 NGATCGCATCTTAAGCTGGTTCATGACCGCTCTTAAGCTGACACCGATGATATTGCT 724
QY 271 IleGluSerLeuHisLysIleGluSerPheValIleGluHisThrAlaAspSerSerAla 290
Db 725 CTCGAGTGCCTCCACAAGATCGAGAGCTTTGTGCTCGAA-----GACACCGCTGTC 775
QY 291 SerAspAlaGlnAlaGluSerGlnIleGlnArgIleGlnThrLeu 305
Db 776 GCCGACTCACAAACGGAATCGCAATCTCAAAGGATGCATACCTTG 820

RESULT 3
LOCUS AK109730 1918 bp mRNA linear PLN 24-JUL-2003
DEFINITION Oryza sativa (japonica cultivar-group) cDNA clone:002-146-B12, full
insert sequence.
ACCESSION AK109730
VERSION AK109730.1 GI:32994939
KEYWORDS FLI_CDNA; oligo capping.
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
REFERENCE 1
AUTHORS The Rice Full-Length cDNA Consortium, National Institute of
Agrobiological Sciences Rice Full-Length cDNA Project Team:,
Kikuchi,S., Satoh,K., Nagata,T., Kawagashira,N., Doi,K.,
Kishimoto,N., Yazaki,J., Ishikawa,M., Yamada,H., Ooka,H., Hotta,I.,
Kojima,K., Namiki,T., Ohneda,E., Yahagi,W., Suzuki,K., Li,C.,
Ohtsuki,K., Shishiki,T., Foundation of Advancement of International
Science Genome Sequencing & Analysis Group:, Ootomo,Y., Murakami,K.,
Iida,Y., Sugano,S., Fujimura,T., Suzuki,Y., Tsunoda,Y.,
Kurosaki,T., Kodama,T., Masuda,H., Kobayashi,M., Xie,Q., Lu,M.,
Narikawa,R., Sugiyama,A., Mizuno,K., Yokomizo,S., Niikura,J.,
Ikeda,R., Ishibiki,J., Kawamata,M., Yoshimura,A., Miura,J.,
Kusumegi,T., Oka,M., Ryu,R., Ueda,M., Matsubara,K., RIKEN,
Kawai,J., Carninci,P., Adachi,J., Aizawa,K., Arakawa,T., Fukuda,S.,
Hara,A., Hashidume,W., Hayatsu,N., Imotani,K., Ishii,Y., Itoh,M.,
Kagawa,I., Kondo,S., Konno,H., Miyazaki,A., Osato,N., Ota,Y.,
Saito,K., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T.,
Yoshino,M. and Hayashizaki,Y.
TITLE Collection, mapping, and annotation of over 28,000 cDNA clones from
japonica rice
JOURNAL Science 301 (5631), 376-379 (2003)
MEDLINE 22752273

```


Company, 1 Innovation Way, P.O. Box 6104, Newark, DE 19714-6104, USA

FEATURES
source

Location/Qualifiers
1. .l323
/organism="Triticum aestivum"
/mol_type="mRNA"
/db_xref="taxon:4565"
/clone="wri.pk0099.b12.fis"

ORIGIN

Alignment Scores:
Pred. No.: 2,34e-57 Length: 1323
Score: 861.50 Matches: 183
Percent Similarity: 78.75% Conservative: 32
Best Local Similarity: 67.03% Mismatches: 52
Query Match: 57.77% Indels: 6
DB: Gaps: 2

US-10-660-226-10 (1-305) x BT009514 (1-1323)

Qy 14 AlaAlaGlyPheGlySerSerArgHisArgGlyGlyLeuGlnAlaProThrGlySerLeu 33
|||||
Db 22 GCGCGAGGATTCCTCCCGCGGACAGTCCCGCGGTGAGGCTCGACCCAGATCCGCG--- 78

Qy 34 ArgValAlaAspProAlaGlyProAla-ValAlaValArgAlaArgGlySerLysProVa 53
|||||
Db 79 -----GCGCGCGCGCGTGTCTTCGCGACCGACGCGCGGAGCGCGAGCATCCCAT 132

Qy 53 1-----AlaProLeuArgLeuArgAlaLysLysSerSerClyGlyHisGluAsnSerHi 71
|||||
Db 133 CGTGGCGCGACGCTCAAGGCCCTGTGTGCGCAAAATCGCGAGGTACTGAGAAAGCCCA 192

Qy 71 sAsnSerValAspGluAlaLeuLeuLysArgLysSerSerGluGluValLeuPheTyrLe 91
|||||
Db 193 CTATTCTGCTGATGAGGCTCTGCTACTTAAGCAAAAGCAGAGGACGTGCTCCCTTACCT 252

Qy 91 uAenglyArgCysIleTyrLeuValGlyMetMetClySerGlyLysSerThrValGlyLy 111
|||||
Db 253 GAATGACCGCTGTGTTTATCTAGTTTGAATGATGGGTTCCGCCAAAATACAGTTGGGA 312

Qy 111 sIleMetSerGluValLeuGlyTyrSerPheAspSerAspLysLeuValGlnAla 131
|||||
Db 313 GATAATAGCTGAAGTACTAGGCTATTCAATCTTTGACAGTGATAAGCTGGTTGAGCAGTC 372

Qy 131 aValGlyMetProSerValAlaGlnIlePheLysValHisSerGluAlaPhePheArgAs 151
:|||||
Db 373 TGTGTGCATACCGTCGGTGGCTGAGATTTTCAGGTCACAGTCAAGCATCTTCTCAGAG 432

Qy 151 pAenGluSerSerValLeuArgAspLeuSerSerMetArgLysValValAlaThrGl 171
|||||
Db 433 TAAAGAGAGTGAGGTACTAAGGATTTGTCGTCAATGACCGCATTAATGTTGCAACAGG 492

Qy 171 yGlyGlyAlaValIleArgProIleAsnTyrArgTyrMetLysArgGlyLeuSerValTr 191
|||||
Db 493 AGGTGTGCGGTGATACGACCAATCAATGGAGTTATATGAAGAAAGAGTCTCACTATTG 552

Qy 191 pLeuAspValProLeuAspAlaLeuAlaArgArgIleAlaLysValGlyThrAlaSerAr 211
|||||
Db 553 GTTAGATGTTCCATTGGACGCCCTTGCAGAAAGAGATGTCGCGGTGGTACTGGGTGACG 612

Qy 211 gProLeuLeuAspGlnProSerGlyAspProTyrAlaMetAlaPheSerLysLeuSerMe 231
|||||
Db 613 ACCCTCTCTGCATCAGGAATCTGGTGATCTTATGCAAGAGGCCCTATGCCAACTTACAGC 672

Qy 231 tLeuAlaGlnGlnArgGlyAspAlaTyrAlaAsnAlaAspValArgValSerLeuGluGl 251
|||||
Db 673 ACTTTTGAACAAGAAGATGGATTCTATATGCTTAATGCTGATGCCGAGTTTCCCTTGA 732

Qy 251 ulleAlaCysLysGlnGlyHisAspAspValSerLysLeuThrProThrAspIleAlaIl 271
|||||
Db 733 TATTGATTTCAACAAAGGACATATGATGTGAATGATCTTACACCAAGTCGATCGCTAT 792

Qy 271 eGluSerLeuHisLysIleGluSerPheValIleGlu 283

```

Db          793 TGAGGCATTCTTAAGATGGAGAGCTTCTTACTGAG 829
||||:|||| |||:|||||:|||||:||||: |||
RESULT 5
AP004089/c
LOCUS      111777 bp      DNA      linear      HTG 21-MAR-2002
DEFINITION Oryza sativa (japonica cultivar-group) chromosome 2 clone
            OJ1308_A10, *** SEQUENCING IN PROGRESS ***,
ACCESSION  AP004089
VERSION     AP004089.1 GI:15281368
KEYWORDS    HTG; HTGS PHASE2.
SOURCE      Oryza sativa (japonica cultivar-group)
ORGANISM    Oryza sativa (japonica cultivar-group)
            Sukariyaka, Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzaceae; Oryza.
REFERENCE   1
AUTHORS     Sasaki, T., Matsumoto, T. and Yamamoto, K.
TITLE       Oryza sativa nipponbare (GA3) genomic DNA, chromosome 2, BAC
            clone:OJ1308_A10
JOURNAL     Published Only in Database (2001)
REFERENCE   2 (bases 1 to 111777)
AUTHORS     Sasaki, T., Matsumoto, T. and Yamamoto, K.
TITLE       Direct Submission
JOURNAL     Submitted (22-AUG-2001) Takuji Sasaki, National Institute of
            Agrobiological Resources, Rice Genome Research Program; Kamondai
            2-1-2, Tsukuba, Ibaraki 305-8602, Japan
            (E-mail:tsasaki@nart.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/,
            Tel:81-298-38-7441, Fax:81-298-38-7468)
COMMENT     The nucleotide sequence of this BAC clone was generated by
            combining Monsanto and RGP-Japan sequencing data.
            NOTE: It currently consists of 1 contigs. Gaps between the contigs
            are represented as runs of N. The order of the pieces is believed
            to be correct as given, however the sizes of the gaps between them
            are based on estimates that have provided by the submitter. This
            sequence will be replaced by the finished sequence as soon as it is
            available and the accession number will be preserved.
            * NOTE: This is a 'working draft' sequence.
            * This sequence will be replaced
            * by the finished sequence as soon as it is available and
            * the accession number will be preserved.
FEATURES             Location/Qualifiers
     source           1..111777
                     /organism="Oryza sativa (japonica cultivar-group)"
                     /mol_type="genomic DNA"
                     /cultivar="Nipponbare"
                     /db_xref="taxon:39947"
                     /chromosome="2"
                     /clone="OJ1308_A10"
ORIGIN
Alignment Scores:
Pred. No.:          9,93e-52          Length:          111777
Score:              836.50            Matches:         263
Percent Similarity: 34.24%            Conservative:    14
Best Local Similarity: 32.51%          Mismatches:     25
Query Match:        54.82%            Indels:         512
DB:                  2                Gaps:           13

US-10-660-226-10 (1-305) x AP004089 (1-111777)

Qy      1 MetGluAlaGlyGlyValGlyLeuAlaLeuGlnAlaArgAlaAlaGlyPhe---GlySer 19
Db      65897 ATGAGGCG---GCGGTGGGGCTGGCGCTCGATGTCGGCGCGCGGGGTTCGGCGGCTCC 6584
Qy      20 SerArgHisArgGlyGlyLeu-----GlnAlaProThrGlySerLeuArg 34
Db      65840 GACCGCGCGGAGCGCGCTCTACGCGCGGAGGGCGCGCGGATCGGGAGCTTCAGG 6578
Qy      35 ValAlaAspProAlaGlyProAlaValAlaValAlaArgGlySerIysProValAla 54
Db      65780 GTGCTGACGCGCGGTGGCGAAGCGCGCTGTGTGGGCTCGCGGGTCCAAAGCGGTGCC 6572

```


Db 63571 CGAGACTTGTCTATCAACATACCGTTGACATCCGTTGGTGACTCCAGCGTCACTC 63512

Qy 297 rGlnileGlnArgileGlnThrLeu 305

Db 63511 ACGTGCTCAGAGATACAGACCTTG 63487

RESULT 6

AP004774

LOCUS

DEFINITION

Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 2,

PAC clone:P0431B06.

ACCESSION

AP004774

VERSION

AP004774.3

KEYWORDS

ORGANISM

Oryza sativa (japonica cultivar-group)

Oryza sativa (japonica cultivar-group)

Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

Ehrhartoideae; Oryzeae; Oryza.

1

Sasaki, T., Matsumoto, T. and Yamamoto, K.

Oryza sativa nipponbare (GA3) genomic DNA, chromosome 2, PAC

clone:P0431B06

Published Only in Database (2002)

2 (bases 1 to 154796)

Sasaki, T., Matsumoto, T. and Yamamoto, K.

Direct Submission

Submitted (20-PEB-2002) Takuji Sasaki, National Institute of

Agrobiological Sciences, Rice Genome Research Program; Kannondai

2-1-2, Tsukuba, Ibaraki 305-8602, Japan

E-mail:tsasaka@nias.affrc.go.jp, URL:htp://rgp.dna.affrc.go.jp/,

Tel:81-298-38-7441, Fax:81-298-38-7468

On Apr 14, 2004 this sequence version replaced gi:33468820.

Genes were predicted from the integrated results of the following:

GENSCAN (<http://CCR-081.mit.edu/GENSCAN.html>), FGENESH

(<http://www.softberry.com/>), GeneMark.hmm

(<http://opal.biology.gatech.edu/GeneMark/>), GlimmerM

(http://www.tigr.org/tdb/glimmer/glmr_form.html), RiceHMM

(<http://rgp.dna.affrc.go.jp/RiceHMM/>), SplicePredictor

(<http://bioinformatics.iastate.edu/cgi-bin/sp.cgi>), sim4

(<http://globin.cse.psu.edu/html/docs/sim4.html>), gap2

(<http://www.tigr.org/software/glimmer/>), BLASTN and BLASTX. The

genomic sequence was searched against NCBI NonRedundant Protein

database, nr (<ftp://ncbi.nlm.nih.gov/blast/db>) and the cDNA

sequence database at RGP or DDBJ. Protein homologs of the coding

regions were searched against NCBI NonRedundant Protein database

with BLASTP. ESTs represent the identified cDNA sequences using

BLASTN with the corresponding DDBJ accession no. and RGP clone ID.

Full-length cDNAs represent the identified cDNA sequences using

BLASTN with identity or significant homology to a protein is

A gene with identity or significant homology to a protein is

classified based on the protein name to indicate the homology level

such as same name, 'putative-' and '-like protein'. A gene without

significant homology to any protein but with full-length cDNA or

EST homology (covering almost the entire length of partial

sequence) is classified as an 'unknown' protein. A gene predicted

by two or more gene prediction programs is classified as a

'hypothetical' protein according to IRGSP standard. A gene

predicted by a single gene prediction program is also classified as a

probable 'hypothetical' protein and is included as a

miscellaneous feature of the sequence.

The orientation of the sequence is from T7 to SP6 of the PAC clone.

This sequence of P0431B06 clone has an overlap with OSJNBa0078N11

(DDBJ: AP005848) clone at 5' end and with OJ1288 G09 (DDBJ:

AP004119) at 3' end. Detailed information on overlap and assembly

quality together with annotation of this entry is available at

<http://rgp.dna.affrc.go.jp/GenomeSeq.html>.

Location/Qualifiers

1..154796

/organism="Oryza sativa (japonica cultivar-group)"

/mol_type="genomic DNA"

gene

misc_feature

gene

misc_feature

gene

misc_feature

gene

misc_feature

gene

misc_feature

gene

misc_feature

gene

misc_feature

gene

misc_feature

gene

misc_feature

gene

misc_feature

gene

misc_feature

gene

misc_feature

gene

misc_feature

gene

misc_feature

gene

misc_feature

gene

misc_feature

gene

misc_feature

gene

misc_feature

gene

misc_feature

gene

misc_feature

gene

misc_feature

gene

misc_feature

gene

misc_feature

gene

misc_feature

gene

misc_feature

gene

misc_feature

gene

misc_feature

gene

misc_feature

gene

misc_feature

gene

misc_feature

gene

misc_feature

gene

misc_feature

gene

misc_feature

gene

misc_feature

gene

misc_feature

gene

misc_feature

gene

misc_feature

gene

/chromosome="2"
/clone="P0431B06"
747..3402
/gene="P0431B06.1"
join(<747..1033,1549..1595,1683..1737,1842..1914,
2001..2088,2218..2300,2401..2421,2445..2821,2969..3033,
3141..33402)
/gene="P0431B06.1"
/note="supported by full-length cDNA(s): AK0566687"
747..3402
/gene="P0431B06.1"
/note="contains full-length cDNA(s): AK066667
non-coding transcript
probably inactive due to 3'UTR too long in CDS"
complement(4342..4808)
/gene="P0431B06.2"
complement(4342..4808)
/gene="P0431B06.2"
/note="supported by full-length cDNA(s): AK058832"
complement(4442..4759)
/gene="P0431B06.2"
/note="contains full-length cDNA(s): AK058832"
/codon_start=1
/product="unknown protein"
/protein_id="BAD15627.1"
/db_xref="GI:46390196"
/translation="MKKAVLVGRGRRREREREGRRRRRPPAVAHTRCPLLPSPHMPATASAG
RIGVGRRLLDPASRPPEGAARCGARRRPPRAVAHTRCPLLPSPHMPATASAG
1E6"
complement(6499..7930)
/gene="P0431B06.3"
complement(join(6499..6558,6932..7039,7135..7255,
7869..7930))
/gene="P0431B06.3"
/note="start and end point are not identified"
complement(join(6499..6558,6932..7039,7135..7255,
7869..7930))
/gene="P0431B06.3"
/note="contains EST(s): AU172742(E20094),C73662(E20094)"
/codon_start=1
/product="dolichol-phosphate mannosyltransferase-like"
/protein_id="BAD15628.1"
/db_xref="GI:46390197"
/translation="MVCWCIERQLVARDPRSLSKHLVQMKHIFKIIAMLVAVSAI
WALLETATVPRSYTLLPIYLVALGCVGLFMVGLMFFPTCPQEAVALQDILEA
KEFLSKGVGVGSR"
join(9079..9119,9183..9349,9987..10000)
/gene="P0431B06.4"
join(9079..9119,9183..9349,9987..10000)
/gene="P0431B06.4"
/note="hypothetical ORF
predicted by GENSCAN
this category is not included in IRGSP standard"
complement(10085..21473)
/gene="P0431B06.5"
complement(join(10085..10456,10633..10695,10796..10912,
11113..11202,11457..11544,12022..12091,12505..12583,
13791..13835,14487..14573,14661..14744,14864..15043,
15122..15262,15433..15541,15971..16026,16437..16484,
16581..16719,17270..17377,17698..17724,17885..17977,
18397..18476,18611..18700,18846..18936,19063..19354,
21171..21473))
/gene="P0431B06.5"
/note="supported by full-length cDNA(s): AK119455"
complement(join(10325..10456,10633..10695,10796..10912,
11113..11202,11457..11544,12022..12091,12505..12583,
13791..13835,14487..14573,14661..14744,14864..15043,
15122..15262,15433..15541,15971..16026,16437..16484,
16581..16719,17270..17377,17698..17724,17885..17977,
18397..18476,18611..18688))
/gene="P0431B06.5"
/note="contains EST(s): AU069794(E3332)
contains full-length cDNA(s): AK119455"

Db 1898 TATTGGTCTTTGATAG-GTTTGTATTTTGGCTATTGTCGAAGAGATCTTGCCCTTGGT 1956
QY 125 - - - - -AspLysLeuValGlu 129
Db 1957 GCTCATGGTAATCTAAGCTTTTATATTCTGGAATTGTTTGCAGTGTATAAATTTGGTCGA 2016
QY 130 GluAlaValGlyMetProSerValAlaGlnIlePheLysValHisSerGluAlaPhePhe 149
Db 2017 CAAGCTGTGGGCATGCCCTTCAGTCGCTCAAAATTTTCAAGGTTTCATAGTGAAGCCTTCTTT 2076
QY 150 AtcAspAsnGlu- - - - - 153
Db 2077 AGGGAATAGGTAATCTTTTCTGAAATCCATCTTCTGTAAACATGATCTACTATCT 2136
QY 153 - - - - - 153
Db 2137 TAGCAATTTGTTACTTGTAGTATATCAGAGAAATCTGAGACTGATTATTTTCTCTGTG 2196
QY 154 - - - - -SerSerValLeuArgAspLeuSerSerMetArgArgLeu 166
Db 2197 GCTCTCTTTTATGCATTCACAGACTAGTGTCTTGAGGGAATTTGTCTCAATGAAGCGATT 2256
QY 167 ValValAlaThrGlyGlyAlaValIleArgProIleAsnTrp- - - - - 181
Db 2257 GTTGTGTACTGGAGGTGGTGTCTGTATCCGACCAGTTAACTGGTATCTGGAGTTCACT 2316
QY 181 - - - - - 181
Db 2317 TCTATTTTCTGTTGGAATAGTTTATTTTGGCTTCTCTATTATTTGATTAAACAATAT 2376
QY 182 - - - - -ArgTyrMetLysArgGlyLeuSerValTrpLeuAs 193
Db 2377 GTGATCTCTTGTGTTATATCAGAAATACATGAAGAGGCTTATCTGTTTGGTTGA 2436
QY 193 pValProLeuAspAlaLeuAlaArgIleAlaLysValGlyThrAlaSerArgProLe 213
Db 2437 TGTGGCCTTGGAGCGCTCTTGTCTAGGCGTATGCTAAAGTGGGAGCTCCCTCCGCTCT 2496
QY 213 uLeuAspGlnProSerGlyAspProTyrAlaMeta- - - - - 225
Db 2497 TCTAGATCAACCATCTGTTGATCCATACACATAGT- GACTACTGGCCCATGAATTTAT 2555
QY 225 - - - - - 225
Db 2556 GCCATGCATCAATATGATATATTTCAGATGTATTTTAAATTTTAAATCAATCAATA 2615
QY 225 - - - - - 225
Db 2616 TAGGCTCATTCATCAATGTAGAAATGAAAACTGCAAGTATGATCCTGCACATCTGCAC 2675
QY 225 - - - - - 225
Db 2676 AACTGGTAGTGAATCTTTCATTTCAGTGTAGTATGAGTGTCTCAGTGAATCTTCTCT 2735
QY 226 - - - - -PheSerLysLeuSerMetLeuAlaGlnArgGlyAspAlaTyrAlaAs 242
Db 2736 ATCTTAGGCTTTTCTTAACTCAGCATGCTCGCGGAGCAAGGGCGATGCTTATGCAA 2795
QY 242 nAlaAspValArgValSerLeuGlu- - - - - 250
Db 2796 TGCTGATGTGAGGTTTCTCTTGAAGGTATACTTCACAAGTTTTCGCAAAATCCGCGCT 2855
QY 250 - - - - - 250
Db 2856 TTCCCAATTTAGAGAATATGTGCTAGGTTATGTGCGGACACATTACTTTTAAAGTAAGTCT 2915
QY 251 - - - - -GluIleAl 253
Db 2916 ATAACCTCCATTTTCCCTGCATAGATAACATATATGTTTTGGAACTTGCAGAGATTGC 2975
QY 253 aCysLysGlnGlyHisAspValSerLysLeuThrProThrAspIleAlaIleGluSe 273

Db 2976 ATCTAAACAGGGTCATGATGATGTATCGAAGCTAAACCACTGACATTTGCTATTGAGGT 3035
QY 273 rLeu- - - - - 274
Db 3036 -ATGCATGATAAGCTATTCTGAACCCACACANTCAGGACATATGATCTTATTTTCT 3094
QY 275 - - - - -HisLysIle 277
Db 3095 AAGTACTGAAACAATTTCCCTAAACAACATTTCTGGCATTTTACAGTCGTTTCATAAGAT 3154
QY 277 eGluSerPheValIleGluHisThrAlaAspSerSerAlaSerAspAlaGlnAlaGluSe 297
Db 3155 CCAGAACCTTTGTCAATGAACATACCCTTGCAATCCGTTGGTGTGACTCCAGGCTGACTC 3214
QY 297 rGlnIleGlnArgIleGlnThrLeu 305
Db 3215 ACGTGTCTCAGAGGATACAGACCTTG 3239
RESULT 7
AP005848 155168 bp DNA linear PLN 15-APR-2004
LOCUS Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 2,
DEFINITION BAC clone:OSJNBa0078N1.
ACCESSION AP005848
VERSION AP005848.3 GI:46390919
KEYWORDS
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Oryza sativa (japonica cultivar-group)
REFERENCE 1 Oryza sativa (japonica cultivar-group)
AUTHORS Oryza sativa (japonica cultivar-group)
TITLE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
JOURNAL Submitted (24-OCT-2002) Takuji Sakai, National Institute of
Agrobiological Sciences, Rice Genome Research Program; Kannondai
2-1-2, Tsukuba, Ibaraki 305-8602, Japan
REFERENCE 2 (bases 1 to 155168)
AUTHORS (E-mail:tsasakia@nias.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/,
Tel:81-298-38-7441, Fax:81-298-38-7468)
TITLE Genes were predicted from the integrated results of the following:
GENSCAN (http://CCR-081.mit.edu/GENSCAN.html), FGENESH
(http://www.softberry.com/), GeneMark.hmm
(http://opal.biology.gatech.edu/GeneMark/), GlimmerM
(http://www.tigr.org/tdb/glimmer/glmr_form.html), RiceHMM
(http://rgp.dna.affrc.go.jp/RiceHMM/), SplicePredictor
(http://bioinformatics.iastate.edu/cgi-bin/sp.cgi), eim4
(http://globin.cse.psu.edu/html/docs/eim4.html), gap2
(http://www.tigr.org/software/glimmer/), BLASTN and BLASTX. The
genomic sequence was searched against NCBI NonRedundant Protein
database, nr (ftp://ncbi.nlm.nih.gov/blast/db) and the cDNA
sequence database at RGP or DBJ. Protein homologs of the coding
regions were searched against NCBI NonRedundant Protein database
with BLASTP. ESTs represent the identified cDNA sequences using
BLASTN with the corresponding DBJ accession no. and RGP clone ID.
Full-length cDNAs represent the identified cDNA sequences using
BLASTN with the corresponding DBJ accession no.
A gene with identity or significant homology to a protein is
classified based on the protein name to indicate the homology level
such as same name, 'putative-' and '-like protein'. A gene without
significant homology to any protein but with full-length cDNA or
EST homology (covering almost the entire length of partial
sequence) is classified as an 'unknown' protein. A gene predicted
by two or more gene prediction programs is classified as a
'hypothetical' protein according to IRGSP standard. A gene
predicted by a single gene prediction program is also classified as
a probable 'hypothetical' protein and is included as a
miscellaneous feature of the sequence.

The orientation of the sequence is from M13rev to -21M13 of the BAC clone. This sequence of OSJNBa0078N11.6 clone has an overlap with OSJNBb0024K03 (DDBJ: AP005733) clone at 5' end and with F0431B06 (DDBJ: AP004774) at 3' end. Detailed information on overlap and assembly quality together with annotation of this entry is available at <http://tgg.dna.affrc.go.jp/GenomeSeq.html>.

FEATURES		Location/Qualifiers
Source		1..155168 /organism="Oryza sativa (japonica cultivar-group)" /mol_type="genomic DNA" /cultivar="Nipponbare" /db_xref="taxon:39947" /chromosomes="2" /clone="OSJNBa0078N11" join(7880..7927,8056..8262) /gene="OSJNBa0078N11.1" join(7880..7927,8056..8262) /genes="OSJNBa0078N11.1" /notes="hypothetical ORF predicted by GENSCAN this category is not included in IRGSP standard" complement(10104..10938) /genes="OSJNBa0078N11.2" complement(join(<10104..10227,10705..10718,10837..>10938)) /genes="OSJNBa0078N11.2" /notes="start and end point are not identified" complement(join(10104..10227,10705..10718,10837..10938)) /genes="OSJNBa0078N11.2" /notes="predicted by GeneMark.hmm etc." /codon_start=1 /product="hypothetical protein" /protein_id="BAD16434.1" /db_xref="GI:46390920" /translation="MTTAMTAGRFGAARHGQRARAGAEAGDGGDQAGSGTGRHDGK GATRPPTARSQWEMGTASRGARAGAAAGEKNA" complement(11100..12051) /genes="OSJNBa0078N11.3" complement(join(<11100..11198,11911..>12051)) /genes="OSJNBa0078N11.3" /notes="start and end point are not identified" complement(join(11100..11198,11911..12051)) /genes="OSJNBa0078N11.3" /codon_start=1 /product="circumsporozoite protein-like" /protein_id="BAD16435.1" /db_xref="GI:46390921" /translation="MRRHAQPHGLFWNARTGGADRGPRDPISAEIAPTWRRLRGCHAG WREGNVGLTDFGRRELALGARVQIGDRAEEG" join(13129..13192,13295..13335,13511..13587,13980..14098, 14169..14218) /genes="OSJNBa0078N11.4" join(13129..13192,13295..13335,13511..13587,13980..14098, 14169..14218) /genes="OSJNBa0078N11.4" /notes="hypothetical ORF predicted by GENSCAN this category is not included in IRGSP standard" complement(15094..15419) /genes="OSJNBa0078N11.5" complement(join(<15094..15102,15246..>15419)) /genes="OSJNBa0078N11.5" /notes="start and end point are not identified" complement(join(15094..15102,15246..15419)) /genes="OSJNBa0078N11.5" /notes="predicted by GENSCAN etc." /codon_start=1 /product="hypothetical protein" /protein_id="BAD16436.1" /db_xref="GI:46390922" /translation="MKSGGPTKDGFTSGLYCPPLPSCGGQAGCGGLWPGGARWRR RAGGWNPAGFCLAVDI" 15790..16489 /genes="OSJNBa0078N11.6"
gene		join(<15790..15967,16434..>16489) /gene="OSJNBa0078N11.6" /note="start and end point are not identified" join(15790..15967,16434..16489) /gene="OSJNBa0078N11.6" /note="predicted by GENSCAN etc." /codon_start=1 /product="hypothetical protein" /protein_id="BAD16437.1" /db_xref="GI:46390923" /translation="MAGRGPGCEAARWQLAAGDVAGGAAGAAHGGFGGEAAWRCAGALA AGGAGDAAGAEVAGLRGGGEGGEGGRRE" complement(18875..19527) /gene="OSJNBa0078N11.7" complement(join(<18875..18932,19289..>19527)) /gene="OSJNBa0078N11.7" /note="start and end point are not identified" complement(join(18875..18932,19289..19527)) /genes="OSJNBa0078N11.7" /note="predicted by GeneMark.hmm etc." /codon_start=1 /product="hypothetical protein" /protein_id="BAD16438.1" /db_xref="GI:46390924" /translation="MPVTGROWRRRGRHCRSTIPARRGGAASRPLLPMLAGRHR CRSTIPSAHREAVNGATLTGSGGGGIGPRWEINGHADCRLKLPQVARGG" complement(join(22369..22410,22804..23458,24470..>24483)) /gene="OSJNBa0078N11.8" complement(join(22369..22410,22804..23458,24470..24483)) /genes="OSJNBa0078N11.8" /note="hypothetical ORF predicted by GENSCAN this category is not included in IRGSP standard" join(28381..28444,28479..28567) /gene="OSJNBa0078N11.9" join(28381..28444,28479..28567) /gene="OSJNBa0078N11.9" /note="hypothetical ORF predicted by GlimmerM this category is not included in IRGSP standard" 28959..29387 /genes="OSJNBa0078N11.10" join(<28959..29126,29274..>29387) /gene="OSJNBa0078N11.10" /note="start and end point are not identified" join(28959..29126,29274..29387) /gene="OSJNBa0078N11.10" /note="predicted by GENSCAN etc." /codon_start=1 /product="hypothetical protein" /protein_id="BAD16439.1" /db_xref="GI:46390925" /translation="MGRWPPSPSEEGARSGQWSAKAGRCDDPAIITNSNRATVLS PPFLIHFRYRPQAGDFADNRSGDDEAGERREEGCGRRWSSVSSWGE" 30927..34707 /gene="OSJNBa0078N11.11" 30927..31672 /gene="OSJNBa0078N11.11" /note="5' terminal repeat" /rpt_type="terminal" 32826..34707 /gene="OSJNBa0078N11.11" /note="probably inactive due to including stop codon(s) in CDS pseudogene, GAG-POL precursor" /pseudo 34780..37821 /gene="OSJNBa0078N11.12" 34780..37821 /gene="OSJNBa0078N11.12" /note="probably inactive due to including stop codon(s) in CDS pseudogene, GAG-POL precursor"
misc_feature		
repeat_unit		
misc_feature		
gene		
misc_feature		

Alignment Scores:		/pseudo	
Pred. No.:	1.45e-51	Length:	155168
Score:	836.50	Matches:	263
Percent Similarity:	34.24%	Conservative:	14
Best Local Similarity:	32.51%	Mismatches:	25
Query Match:	54.82%	Indels:	512
DB:	8	Gaps:	13
US-10-660-226-10 (1-305) x AP005848 (1-155168)			
QY	1	MetGluAlaGlyValGlyValGlyLeuAlaLeuGlnAlaArgAlaGlyPhe---	GlySer 19
Db	145666	ATGGAGGCG---GGCGTGGGCTGGCGCTGCAGTCGGCGCGGGGGTTCGGCGGCTCC	145722
QY	20	SerArgHisArgGlyGlyLeu-----GlnAlaProThrGlySerLeuArg 34	
Db	145723	GACCGCGCGGAGCGGCTCTACGGCGGCGAGGGCGGCGCGGATCGGGAGCTTGAGG	145782
QY	35	ValAlaAspProAlaGlyProAlaValAlaValArgAlaArgGlySerLysProValAla 54	
Db	145783	GTCGCTGAGCGCGGGTGGCGAGCGCGTGTGTGGGCTCGCGGTCGAAGCGGTCGCC	145842
QY	55	ProLeuArgLeuArgAlaLysLysSerSerGly-----	65
Db	145843	CCG-----CTCGGTGCCAAGAAATCGTCCGG-AGGTACAGGGTCGCTCTCTTCTT	145895
QY	65	-----	65
Db	145896	CGTTCTGTCAAGATTGGTTGATTTGGCAAAAGTTTGTGTTTTTTTGGACAACTGGCG	145955
QY	65	-----	65
Db	145956	GTCTGGCGCCATTGGGCCCTTTTGTGTGTCCGATTTTGTGTCGATCGGAGTTGTG	146015
QY	65	-----	65
Db	146016	ATATTTGGAGGAACACATGGAACGGGTGCGTTGATTCGGTCAAAAAAAGAGCTT	146075
QY	65	-----	65
Db	146076	GCAAGTAAGGAATTAGGTGTAAGGCTGGCTTCGAATGGATCGAATCTGCAAAAGTTGT	146135
QY	65	-----	65
Db	146136	GCCTCTTTCTGTTGTTGCTTTTGTATGCGGTCCATTTTGGTTCAGATTTTCAGTCCTA	146195
QY	65	-----	65
Db	146196	TTGTTCCGATTAGATTGATCAAAATGGATAGGTTGATTTGGGATTAATTATTTGGAAGGTAG	146255
QY	65	-----	65
Db	146256	GAAGAGATGAAGTTTCGCCCTATGATTGGTCCCTTCATAGTCAAGCCAACTCTGAAATTC	146315
QY	65	-----	65
Db	146316	ACGAGGTGATTTATGTCTTTCACAAATTTACACAAATTTGACGCTCCTCGTTATATTGT	146375
QY	66	-----GlyHisGluAsnSerHisAsnSerValAspGluAlaLeuLeuLys---	81
Db	146376	TCCTGTGCAGGTTCATGAACATTGCATACTACCTCGGTTGATGAAGCCCTCTTGTCAAAGGTA	146435
QY	81	-----	81
Db	146436	CTGCATAATCAAAATCAAAAGCGACACTGGAGTTTGTGTTTCTTCTCCATCACTCATGT	146495
QY	82	-----ArgLysSerGluGluValLeuPheTyrLeuAsnGly 93	
Db	146496	TCTAACCACTGTGTTATTTTCAGAGAAATACAGAAAGTCTCTCTCTATTGGAATGGA	146555
QY	94	ArgCysIleTyr-----	97
Db	146556	CGGTGTAATTTACCTAGTTGGTTAGTTCTGAAACTTCTCATTTCTCTCCAAATTATCAGCACT	146615
QY	98	-----Leu 98	
Db	146616	GAAAAAGATCAATATATATAGATGGCATCTGAATTTCTGGTTGTATAAATTAAACATCTCTG	146675
QY	99	ValGlyMetMetGlySerGlyLysSerThrValGlyLysIleMetSerGluValLeuGly 118	
Db	146676	CA-GGAATGATGGTTCTCGAAAAAGTACTGTGGGAAAGATCATGTCTGAAGTTTGGGT	146734
QY	119	TyrSerPhePheAspSer-----	124
Db	146735	TATTCGTTCTTTGATAG-GTTTGTATTTTGTCTATTGTCACAGGAGATCTTGCCCTTGGT	146793
QY	125	-----AspLysLeuValGlu 129	
Db	146794	GCTCATGGTAATCTAAGCTTTTATATTTCTGGAAATGTTTGGAGTGAATATTTGGTCGAA	146853
QY	130	GlnAlaValGlyMetProSerValAlaGlnIlePheLysValHisSerGluAlaPhePhe 149	
Db	146854	CAAGCTGTGGCATGCTTCAGTCGCTCAAAATTTTCAAGGTTTCATAGTCAAGCCTTCTTT	146913
QY	150	ArgAspAsnGlu-----	153
Db	146914	AGGGATAATGAGGTAATCTTTTTTCTGAAATCCATCTTCTGCTAACATGATCTACTATCT	146973
QY	153	-----	153
Db	146974	TAGCATTTGTTTACTTTGTTAGTTATCAGAAAGGAAAACTCGAGACTGATTAATTTTCTCTGTG	147033
QY	154	-----SerSerValLeuArgAspLeuSerSerMetArgArgLeu 166	
Db	147034	GCTCTCTTTTATGCATTACAGAGTAGTGTCTTTGAGGATTTTGTCTCAATGAAGCGATTA	147093
QY	167	ValValAlaThrGlyGlyAlaValIleArgProIleAsnTrp-----	181
Db	147094	GTTGTGCTACTGGAGGTGTGCTGTATCCGACCAGTTAACTGGTATCTGGAGTTCACT	147153
QY	181	-----	181
Db	147154	TCATATTTTCTGTTGGAAATAGTTTAATTTTGTGGCTTCTCTATTTTGTATTAACAATATT	147213
QY	182	-----ArgTyrMetLysArgGlyLeuSerValTrpLeuAs 193	
Db	147214	GTGATCTCTTCTGTTGTTATTCAGGAATACATGAAGAGGGCTATCTGTTTGGTTGA	147273
QY	193	pValProLeuAspAlaLeuAlaArgArgIleAlaLysValGlyThrAlaSerArgProLe 213	
Db	147274	TGTGCCCTTGGACGCTCTTGTCTAGGCGTATTGCTAAAGTGGGAGTGCCTCCGCTCTCT	147333
QY	213	uLeuAspGlnProSerGlyAspProTyrAlaMetAla-----	225
Db	147334	TCTAGATCAACCATCTGGTGATCCATACACAATGGT-GACTACTGGCCCCCATGAATTTAT	147392
QY	225	-----	225
Db	147393	GCCATGCATGAATATGATATTTATTTTCAGATGTATTTTAAAAATTTTAAATCAGTCAATA	147452
QY	225	-----	225
Db	147453	TAGGCTCATTTCAATGTAGAAATGAAAAACTGCAAGTATGATCTCTGCACATCTGCAC	147512
QY	225	-----	225
Db	147513	AACTGGTAGTGAATTTCTTTCAATTTTCAGTGTAGTATGATGCTCAGTGGATTTCTCTCT	147572
QY	226	-----PheSerLysLeuSerMetLeuAlaGlnGlnArgGlyAspAlaTyrAlaAs 242	
Db	147573	ATCTTAGGCTTTTCTAACTCAGCATGCTCCGGGAGCAAGGGGGGATGCTTTATGCANA	147632
QY	242	nAlaAspValArgValSerLeuGlu-----	250

Db	147633	TGCTGATGTGAGGGTTTCTCTTGAAGGTATAACTTCACAAGTTTTTCGCAAAATCCGCGCT	147692
QY	250	-----	250
Db	147693	TTCCCAATTAGAGAAATATGTGCTAGGTTATGTGCGACACATTACTTTTAAGTAATGCTTG	147752
QY	251	-----GluIleAl	253
Db	147753	ATAACTCCATTTTCCCTCGCATAGATAACATATATGTTTGGAACTTGCAGAGATTGC	147812
QY	253	aCysLysGlnGlyHisAspValSerLysLeuThrProThrAspIleAlaIleGluSe	273
Db	147813	ATCTAAACAGGTCATGATGATGATCGAAGCTAACACCACTGACATTGCTATTGAGGT	147872
QY	273	rIeu-----	274
Db	147873	-ATGCATGATAAGCTATTCTGAACCCACAATCAGGAGCAATATGAATCTTATTTTCT	147931
QY	275	-----HisLysIle	277
Db	147932	ACGCTACTGAACAATTTCCCTAACAAACATCTTGGCATTTTACAGTCGTTTCATAGAT	147991
QY	277	eGluSerPheValIleGluHisThrAlaAspSerSerAlaSerAspAlaGlnAlaGluSe	297
Db	147992	CGAGAAGCTTTGTCATTGAACATACCGTTGCACATCGGTTGGTGACTCCACAGGCTGACTC	148051
QY	297	rGlnIleGlnArgIleGlnThrIeu	305
Db	148052	ACGTGCTCAGAGCATACAGACCTTG	148076
RESULT 8			
LOCUS	AP003513/c		
DEFINITION	Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 6, PAC clone:P0690H04.		
ACCESSION	AP003513		
VERSION	AP003513.2	GI:42415338	
KEYWORDS	Oryza sativa (japonica cultivar-group)		
SOURCE	Oryza sativa (japonica cultivar-group)		
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.		
REFERENCE	1		
AUTHORS	Sasaki,T., Matsumoto,T. and Yamamoto,K.		
TITLE	Oryza sativa nipponbare(GA3) genomic DNA, chromosome 6, PAC clone:P0690H04		
REFERENCE	2 (bases 1 to 137319)		
AUTHORS	Sasaki,T., Matsumoto,T. and Yamamoto,K.		
JOURNAL	Direct Submission		
AUTHORS	Submitted (11-APR-2001) Takuji Sasaki, National Institute of Agricultural Sciences, Rice Genome Research Program; Kannondai 2-1-2, Tsukuba, Ibaraki 305-8602, Japan		
JOURNAL	(E-mail:tsasaki@nla.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/, Tel:81-298-38-7441, Fax:81-298-38-7468)		
COMMENT	On Feb 5, 2004 this sequence version replaced gi:13603473. Genes were predicted from the integrated results of the following: GENSCAN (http://CCR-081.mit.edu/GENSCAN.html), FGENESH (http://www.softberry.com/), GeneMark.hmm (http://opal.biology.gatech.edu/GeneMark/), GlimmerM (http://www.tigr.org/tdb/glimmerm/glmr_form.html), RiceHMM (http://rgp.dna.affrc.go.jp/RiceHMM/), SplicePredictor (http://bioinformatics.iastate.edu/cgi-bin/sp.cgi), sim4 (http://globin.cse.psu.edu/html/docs/sim4.html), gap2 (http://www.tigr.org/software/glimmer/), BLASTN and BLASTX. The genomic sequence was searched against NCBI NonRedundant Protein database, nr (ftp://ncbi.nlm.nih.gov/blast/db) and the cDNA sequence database at RGP or DDBJ. Protein homologies of the coding regions were searched against NCBI Nonredundant protein database with BLASTP. ESTs represent the identified cDNA sequences using BLASTN with the corresponding DDBJ accession no. and RGP clone ID. Full-length cDNAs represent the identified cDNA sequences using		

BLASTN with the corresponding DDBJ accession no.	
A gene with identity or significant homology to a protein is classified based on the protein name to indicate the homology level such as same name, 'putative-' and '-like protein'. A gene without significant homology to any protein but with full-length cDNA or EST homology (covering almost the entire length of partial sequence) is classified as an 'unknown' protein. A gene predicted by two or more gene prediction programs is classified as a 'hypothetical' protein according to IRGSP standard. A gene predicted by a single gene prediction program is also classified as a probable 'hypothetical' protein and is included as a miscellaneous feature of the sequence.	
The orientation of the sequence is from SP6 to T7 of the PAC clone. This sequence of P0690H04 clone has an overlap with P063BH11 clone (DDBJ: AP005545) at 5' end and an overlap with P0425F05 (DDBJ: AP003569) at 3' end. Detailed information on overlap and assembly quality together with annotation of this entry is available at http://rgp.dna.affrc.go.jp/GenomeSeq.html .	
FEATURES	Location/Qualifiers
1..137319	/organism="Oryza sativa (japonica cultivar-group)"
	/mol_type="genomic DNA"
	/cultivar="Nipponbare"
	/db_xref="taxon:39947"
	/chromosome="6"
	/clone="P0690H04"
2403..6152	/gene="P0690H04.1"
Join(<2403..2534,2937..3068,5832..>6152)	/gene="P0690H04.1"
	/note="start and end point are not identified"
Join(2403..2534,2937..3068,5832..6152)	/gene="P0690H04.1"
	/codon_start=1
	/product="small GTP-binding protein"
	/protein_id="BAD37285.1"
	/db_xref="GI:51535001"
	/translation="MSFLDMFYDVLASIGLWQEKILFLGLDNAGKTLFLPYLMSOE NLAVHQPTQHTPSELSIGRIRFKAPDVGCHRIARVRVDVYAQVDVVVVDADRC RFAESKMDLALLSDLAGLAPFLVGLGKIDIPYAVPEQLCYLGLTLTGKGNVN LAGTVRVEVFMCSSVVRMGYGDGRFWRMSQYIK"
	complement(6476..8173)
	/gene="P0690H04.2"
	complement(6476..8173)
	/gene="P0690H04.2"
	/note="supported by full-length cDNA(s) : AK071845"
	complement(6801..8015)
	/gene="P0690H04.2"
	/note="contains full-length cDNA(s) : AK071845"
	/codon_start=1
	/product="mitochondrial transcription termination factor-like"
	/protein_id="BAD37286.1"
	/db_xref="GI:51535002"
	/translation="MLHQKHLILLSPRATASTLLSRHHGLFSLTRFSAAAAAK SAGHFAVEYLVAATCHLPDQATKASKISHLKSPSPDPAVVAVLAGLGSADTAA VAYDPRLLCAVDPTLAPRLAELAGLSPSQIARLVDPARFRPTVISKLOYVVP LFGSFETLLQALKNNLSLSDLEKVPKPNVALLREGLGACDIKLCIPRLATTS PERSVDMAQAEENVGRGKMRFRHAILAVATISEKTAAGQFLMKTLKWSDAERI AVSKLPVLRSEDSKLSRVSEFLSEVLEGPAYIAYRPAMLTYSLERLMPRHCVLKY LKNDGLIESDSYYSAGVQTEVFVEKYISPVEDTAPHLAEDYAAVSSVKIPTRFLK GPKTGHASQAQA"
	9570..11016
	/gene="P0690H04.3"
	9570..10944
	/gene="P0690H04.3"
	/note="supported by full-length cDNA(s) : AK104692"
	9586..11016
	/gene="P0690H04.3"
	/note="supported by full-length cDNA(s) : AK067095"
	9615..10796
	/gene="P0690H04.3"
	/note="contains EST(s) :

C22491 (E0081), D24730 (R2451), AU173369 (R2451), C22490 (E0081)
 contains full-length cDNA (s): AK104692, AK067095"
 /codon_start=1
 /product="mitochondrial transcription termination
 factor-like"
 /protein_id="BAD37287.1"
 /db_xref="GI:51535003"
 /translation="MVHLRHHMISLLRPAYPHPTAATISPLSSRLRLLESTTAAPV
 SEPAPAVESLWGLTGAQKASKKLSHRSKSPKPAVLAFISDLGLPRKTAAY
 ATADRFECADYENLARVDELGLGLRSOIARLVLATCFSSVGNLGMQLQ
 IVGSPDKILKALRMSSLLGSLEKVKENLELLKQCGMSDFATSPPLYSRLFPANP
 IYRDARVBEGLDRSRMPFRHGLIAVFTSKSVARKIQVMEELGFRDELLEMLII
 RKAPQVASEEKIRQAEFLKREGLRYIAHFRVFLYLSLERELLPFRHLLVLR
 MKGLLDCELDYNTAASERKFRKVPDYKCHIPCLADAYTSSCAGETANGVASLLG
 V"
 complement (join(111360..11924,12422..12816,12949..13290,
 13368..13505,13802..13878,13961..14032,14104..14175,
 14308..14451,14834..14905,14990..15122,15245..15747))
 /gene="P0690H04.4"
 complement (join(111360..11924,12422..12816,12949..13290,
 13368..13505,13802..13878,13961..14032,14104..14175,
 14308..14451,14834..14905,14990..15122,15245..15408))
 /gene="P0690H04.4"
 /notes="supported by full-length cDNA (s): AK065224"
 complement (join(11435..11924,12422..12816,12949..13290,
 13368..13505,13802..13878,13961..14032,14104..14175,
 14308..14451,14834..14905,14990..15122,15245..15747))
 /gene="P0690H04.4"
 /notes="supported by full-length cDNA (s): AK100258"
 complement (join(11598..11924,12422..12816,12949..13290,
 13368..13505,13802..13878,13961..14032,14104..14175,
 14308..14451,14834..14905,14990..15122,15245..15323))
 /gene="P0690H04.4"
 /notes="contains full-length cDNA (s): AK065224, AK100258"
 /codon_start=1
 /product="putative benzothiadiazole-induced somatic
 embryogenesis receptor kinase 1"
 /protein_id="BAD37288.1"
 /db_xref="GI:51535004"
 /translation="WAARRRLLLVLLLCRLAAVLPTSEVALQGFAGAGGNAAF
 QSWDASAPNPTWFHVTCGPGNQVIRLDLGNQSLSGELKPDVWQALQSLEYNSI
 SKIPSELRASLQTLIDLANFTGPIEPLNEIGLSKJLNRLNNNSLQALPMSLT
 IONLEVDLHNNLGGIIPTNGSFHPTPIFSNNPRTFANSDSPSNNGAAVPSGR
 SSASGITTAGAAGAAFLFAPIVTFAMWRWRPHQDFDLLEETPEVHLGQLR
 FTIRELQVATNFSTNLGRGQVKVYGRLLDGLSLAIKRLNDRIGTGRQFLME
 VRIISAHQNLRLQGYCMTPTERLVVYVYMNKSLRTRKRECSDSQOPLDMPTRK
 IALSGRSGISYHEGCDPKLIHRDVYKAAIILDEKLEAVVGFGLARIMDYKSHVT
 GVMGTGHPMBYLTAGRTSKDITVFGYGMIFELISGRKGRFDLVGLANEENARVHDV
 VKKLLEEDRLVLDPNLEIYNGGEGQVREEMRLVQIALICTQESAPSRPRMSTVV
 TMLDEGIASHWDAMORKTIVQASLQGGQGVSEARNDNVANLPDPTLSQPR"
 15552..15773
 /gene="P0690H04.5"
 <15552..>15773
 /gene="P0690H04.5"
 /notes="start and end point are not identified"
 15552..15773
 /gene="P0690H04.5"
 /note="predicted by GlimmerM etc."
 /codon_start=1
 /product="hypothetical protein"
 /protein_id="BAD37289.1"
 /db_xref="GI:51535005"
 /translation="MVRAVAIAVAVGVGSPSGGRLDLRLRQRWPPAAGAPSSRR
 GAGGGQSKQVQLKCLGANRKNPLVSEI"
 complement (join(16607..16955,17062..17153,17446..17579,
 17652..17702,18592..18904))
 /gene="P0690H04.6-1"
 complement (join(16607..16955,17062..17153,17446..17579,
 17652..17702,18592..18904))
 /gene="P0690H04.6-1"
 /notes="supported by full-length cDNA (s): AK109268"
 complement (join(16771..16955,17062..17153,17446..17579,
 17652..17702,18592..18867))

/gene="P0690H04.6-1"
 /note="contains full-length cDNA (s): AK109268
 similar to Arabidopsis thaliana chromosome5, At5g09995"
 /codon_start=1
 /product="unknown protein"
 /protein_id="BAD37290.1"
 /db_xref="GI:51535006"
 /translation="MPSALPPRHTQAPAPPAASASACLRRLRRGLRVSLCYCRAST
 EPLVFPASPSLLRSVAVPSPSRPHANRPGDGGLLVVTVAASAVVLSACFVFL
 SAMRSMLECKAAESLEKSFGRSEKLPETWASVKLVGREICDLAVDLNLSQEURKQ
 VOSKSVVHRAADAOHLHQTTSAPQGNORVTSNRKAAAGEPLLASTVRELRLIAELHS
 GFGVAVSIAGLTWASNFVSKRPNRS"
 complement (18161..19000)
 gene
 Alignment Scores:
 Pred. No.: 8,066-44 Length: 137319
 Score: 733.00 Matches: 205
 Percent Similarity: 38.29% Conservative: 14
 Best Local Similarity: 35.84% Mismatches: 21
 Query Match: 48.03% Indels: 333
 DB: 8 Gaps: 8
 US-10-660-226-10 (1-305) x AP003513 (1-137319)
 QY 66 GlyHisGluAenSerHisAenSerValAspGluAlaLeuLeuLeuLys----- 81
 Db 24116 GSCCAGCAGAGCTTCATAACTCAGTTGATGAAGCCCTCCTGTTAAAGTAAATAACGCT 24057
 QY 81 ----- 81
 Db 24056 TTACAAAATTCCTACTATTAGTTAAATCAAGAGTAAATATGGCTAGAACTTTGTTCT 23997
 QY 82 -----ArgLysSer 84
 Db 23996 TCATCTTTGGCTTCGGCCATGGTATAATGTTTCTAATGTAATTCAGAGAGAAATCA 23937
 QY 85 GluGluValLeuPheTyrLeuAenGlyArgCysIleTyrLeu----- 98
 Db 23936 GAAGAAGTCTTACTTCTACTTGAACGGAGGTGTTATTATTAGTTGGTTAGTTCTGAACC 23877
 QY 98 ----- 98
 Db 23876 TGTATCATTTAATGTTAACTGTTAACTGGAGACTAAATACCAAGTTAAATAACTGA 23817
 QY 99 -----ValGlyMetMetGlySerG1 105
 Db 23816 TCTTATTTGAAATGAATTAATAATATTAGTACGTATTGTAGGAATGATGGCTCAGG 23757
 QY 105 yLysSerThrValGlyLysIleMetSerGluValLeuGlyTyrSerPhePheAspSerAs 125
 Db 23756 GAAGAGTACAGTTGCAAGATATTAGCCGAAGTTTGGGTTATTTCATTCTTTGACAGTGA 23697
 QY 125 pLysLeuValGluGlnAlaValGlyMetProSerValAlaGlnIlePheLysValHisE 145
 Db 23696 TAAAGTTGGTTGAACAGCAGTTGGAATGCCCTCTGTCGCCCAATATTTCACAGGACAG 23637
 QY 145 tGluAlaPhePheArgAspAenGlu----- 153
 Db 23636 TGAACGGTTTTTCAGAGATAATAGGCAATGTTCTTTCTGAAATATGTTCCCTGTTG 23577
 QY 153 ----- 153
 Db 23576 ACAACATGACACTGTTATTATTTTCAGATAGTTTCTATTATTTTGTAGTAGTATTGAATGGAA 23517
 QY 153 ----- 153
 Db 23516 AGATTGGAGTAAACAATTTCTCAGACATGATAGACTAAACAATTTCCATGGTTCCCTGCT 23457
 QY 154 -----SerSerValLeuArgAspLeuSerSerMetArgArgLeuValValAla 170
 Db 23456 TTGCTTAACAGAGTAGTGTCTTGAGGAGCTGTCTCTCAATGCGACAGATTAGTTGTCGA 23397
 QY 170 hrGlyGlyAlaValIleArgProfileAsnTrp----- 181

```
|||||
23396 CTGAGGTGGTCCGTTATTCGGCCAGTTAACTGGTATTGGGATTAATCTGATCTCACGC 23337
QY 181 -----
23336 CCATCCTCTTGTAAATTTGGCTTTTCTGTTTATGGATAATAAGTGAAAGCTCTTGTTGTT 23277
QY 182 -----ArgTyrMetLysArgGlyLeuSerValTrrLeuAspValProLeuAspAlaL 199
23276 TTTTCAGGAATATATGAAGAAGGTCTATCTGTCTGGCTGGATGTCCTTTTGGATGCAC 23217
QY 199 euAlaArgArgIleAlaLysValGlyThrAlaSerArgProLeuLeuAspGlnProSerG 219
23216 TTGCAGGGGTATTGCTCAAGTGGTACTGCCTCTGCTCTCTTCTTGGATCAGCCATCGA 23157
QY 219 lYAspProTyrAla-----
23156 GTGATCCATACAC-AGCGGTAGTTACTCTCTATCTTGAACCTAATCGTTGTGATCATTC 23098
QY 223 -----
23097 CAAGACAGATGCTTTCTTAACCTAGAAGATAACAGTCTTCATGTGAAACACCAATCAGCG 23038
QY 223 -----
23037 TTAGTCTTATAGACCTTGTAGCTTTTGAACCTTGCATGTATTTCATTACCATCTTCATT 22978
QY 224 -----M 224
22977 TTGCTCTTTCGCATGTTGTGCTCATTAATAATGGATGGCTCATGGAATGTTTGTTCGCC 22918
QY 224 etAlaPheSerLysLeuSerMetLeuAlaGlnGlnArgGlyAspAlaTyrAlaAsnAlaA 244
22917 AGGCCCTCTCAAAACTCAGATGCTTGCAGAGCAAGAGCGCATGCTTATGCAAAATGCTG 22858
QY 244 spValArgValSerLeuGlu-----
22857 ATGCTCGAGTTCTCTTGAAGGTATGTTCTTCTTGCATGTCCTTCATCTTAATATGC 22798
QY 250 -----
22797 TCCTAGTTTGTCTGATAGCAGATGGATATCTAGTGTGCTTCAGCGAAGCTTTTCATGTAT 22738
QY 251 -----GlutIleAlaCylLeuGlnG 257
22737 CCACAGAGATTAATACAGTCTGCATATTATTTACCTCTGCAGAAATTCGAGCTAAACAAG 22678
QY 257 lYHisAspAspValSerLysLeuThrProThrAspIleAlaIleGluSerLeuHis---- 275
22677 GCCATCAGCATGTCTCCAAACTAACACACCGCAGCATATTGCTATTGAGGTATTGCACATT 22618
QY 275 -----
22617 CGGGACAATATTTCATGTAAACTTGTTCATTCTCTCTTTGTAAGGACGCTTCATACT 22558
QY 276 -----Lys 276
22557 CTGCCCGATTGAATTTTTTTCCTAACATACATCCTGTTGCATTCCTCAGGCCTTACTCAAG 22498
QY 277 IleGluSerPheValIleGluHis---ThrAlaAspSerSerAlaSerAspAlaGlnAla 295
22497 ATTGAGAATTTCTGCACCGAGCATTCACATCAAGTGGCCCGCTTGGGGACTTAATAGTT 22438
QY 296 GluSerGlnIleGlnArgIleGlnThrLeu 305
22437 GATTCCCAAGATCGAAGGACAAAGGCGCTTG 22408
RESULT 9
AK067948 2127 bp mRNA linear PLN 24-JUL-2003
LOCUS Oryza sativa (japonica cultivar-group) cDNA clone:J013123F09, full
DEFINITION insert sequence.
ACCESSION AK067948
```

```
AK067948.1 GI:32977966
FLI_CDNA; CAP trapper.
Oryza sativa (japonica cultivar-group)
Oryza sativa (japonica cultivar-group)
Euryza sativa (japonica cultivar-group)
Spermatophyta; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Ehrhartoideae; Magnoliophyta; Liliopsida; Poales; Poaceae;
Oryza.
1
The Rice Full-Length cDNA Consortium, National Institute of
Agrobiological Sciences Rice Full-Length cDNA Project Team:
Kikuchi,S., Satoh,K., Nagata,T., Kawagashira,N., Doi,K.,
Kishimoto,N., Yazaki,J., Ishikawa,M., Yamada,H., Ooka,H., Hotta,I.,
Kojima,K., Namiki,T., Ohneda,E., Yahagi,W., Suzuki,K., Li,C.,
Ohtsuki,K., Shishiki,T., Foundation of Advancement of International
Science Genome Sequencing & Analysis Group, Ootomo,Y., Murakami,K.,
Iida,Y., Sugano,S., Fujimura,T., Suzuki,Y., Tsunoda,Y.,
Kurosaki,T., Kodama,T., Masuda,H., Kobayashi,M., Xie,Q., Lu,M.,
Narikawa,R., Sugiyama,A., Mizuno,K., Yokomizo,S., Niikura,J.,
Ikeda,K., Ishibiki,J., Kawamata,M., Yoshimura,A., Miura,J.,
Kusumegi,T., Oka,M., Ryu,R., Ueda,M., Matsubara,K., RIKEN,
Kawai,J., Carninci,P., Adachi,J., Aizawa,K., Arakawa,T., Fukuda,S.,
Hara,A., Hashidume,W., Hayatsu,N., Imotani,K., Ishii,Y., Itoh,M.,
Kagawa,I., Kondo,S., Konno,H., Miyazaki,A., Osato,N., Ota,Y.,
Saito,R., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T.,
Yoshino,M. and Hayashizaki,Y.
Collection, mapping, and annotation of over 28,000 cDNA clones from
japonica rice
Science 301 (5631), 376-379 (2003)
22752273
12869764
2 (bases 1 to 2127)
Adjuchi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Doi,K.,
Fujimura,T., Fukuda,S., Hanagaki,T., Hara,A., Hashizume,W.,
Hayashida,K., Hayashizaki,Y., Hayatsu,N., Hiramoto,K., Hiraoka,T.,
Hori,F., Hotta,I., Iida,J., Iida,Y., Ikeda,R., Imamura,K.,
Imotani,K., Ishibiki,J., Ishii,Y., Ishikawa,M., Itoh,M., Kagawa,I.,
Kanagawa,S., Kato,H., Kawagashira,N., Kawai,J., Kawamata,M.,
Kikuchi,S., Kishikawa-Hirozane,T., Kishimoto,N., Kobayashi,M.,
Kodama,T., Kojima,K., Kurohara,C., Kusumegi,T., Kusunagi,T., Li,C., Lu,M.,
Koya,S., Kurihara,C., Kurosaki,T., Kusumegi,T., L.C., Lu,M.,
Masuda,H., Matsubara,K., Matsuyama,T., Miura,J., Miyazaki,A.,
Mizuno,K., Murakami,K., Murata,M., Nagata,T., Nakamura,M.,
Namiki,T., Narikawa,R., Niikura,J., Nishi,K., Nomura,K.,
Numasaki,R., Ohneda,E., Ohno,M., Ohtsuki,K., Oka,M., Ooka,H.,
Osato,N., Ota,Y., Ootomo,Y., Ryu,R., Saitoh,H., Sakai,C., Sakai,K.,
Sakazume,N., Sano,H., Sasaki,D., Sato,K., Satoh,K., Shibata,K.,
Shinagawa,A., Shiraki,T., Shishiki,T., Sogabe,Y., Sugano,S.,
Sugiyama,A., Suzuki,K., Suzuki,Y., Tagami,M., Tagami-Takeda,Y.,
Tagawa,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A.,
Tova,T., Tsunoda,Y., Ueda,M., Waki,K., Xie,Q., Yahagi,W.,
Yamada,H., Yamamoto,M., Yasunishi,A., Yazaki,J., Yokomizo,S. and
Yoshimura,A.
Direct Submission
Submitted (05-DEC-2001) Shoshi Kikuchi, National Institute of
Agrobiological Sciences, Department of Molecular Genetics, Head of
Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki
305-8602, Japan [E-mail:skikuchi@nias.affrc.go.jp]
Tel:81-29-838-7007, Fax:81-29-838-7007
This clone is one of the 28K full-length cDNA clones from japonica
rice.
URL : http://cdna01.dna.affrc.go.jp/cDNA/
NIAS Rice Full-Length cDNA Project Team: Kikuchi,S., Satoh,K.,
Nagata,T., Kawagashira,N., Doi,K., Kishimoto,N., Yazaki,J.,
Ishikawa,M., Yamada,H., Ooka,H., Hotta,I., Kojima,K., Namiki,T.,
Ohneda,E., Yahagi,W., Suzuki,K., Li,C., Ohtsuki,K., Shishiki,T. and
Yamamoto,M.
FAIS Genome Sequencing & Analysis Group: Ootomo,Y., Iida,Y.,
Fujimura,T., Ikeda,R., Ishibiki,J., Kawamata,M., Kobayashi,M.,
Kodama,T., Kurosaki,T., Kusumegi,T., Lu,M., Masuda,H., Miura,J.,
Mizuno,K., Narikawa,R., Niikura,J., Oka,M., Ryu,R., Sugano,S.,
Sugiyama,A., Suzuki,Y., Tsunoda,Y., Ueda,M., Xie,Q., Yokomizo,S.,
Yoshimura,A., Matsubara,K. and Murakami,K.
Genome Exploration Research Group in Riken Genomic Sciences Center
```


Masuda, H., Matsuura, K., Matsuyama, T., Miura, J., Miyazaki, A., Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakamura, M., Namiki, T., Nariikawa, R., Niikura, J., Nishi, K., Nomura, K., Numasaki, R., Ohneda, E., Ohno, M., Ohtsuki, K., Oka, M., Ooka, H., Osato, N., Ota, Y., Otomo, Y., Ryu, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K., Shibata, K., Shinagawa, A., Shiraki, T., Shishiki, T., Sogabe, Y., Sugano, S., Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, F., Tomaru, A., Toyota, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yahagi, W., Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S. and Yoshimura, A.

TITLE
Direct Submission

JOURNAL
Submitted (05-DEC-2001) Shoshi Kikuchi, National Institute of Agrobiological Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan (E-mail: skikuchi@nias.affrc.go.jp, Tel:81-29-838-7007, Fax:81-29-838-7007)

COMMENT
This clone is one of the 28K full-length cDNA clones from japonica rice. URL : <http://cdna01.dna.affrc.go.jp/cDNA/>
NIAS Rice Full-Length cDNA Project Team: Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T. and Yamamoto, M.

FAIS Genome Sequencing & Analysis Group: Otomo, Y., Iida, Y., Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M., Kodama, T., Kurosaki, T., Kusumegi, T., Lu, M., Masuda, H., Miura, J., Mizuno, K., Nariikawa, R., Niikura, J., Oka, M., Ryu, R., Sugano, S., Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S., Yoshimura, A., Matsuura, K. and Murakami, K.

Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hirooka, T., Hori, F., Iida, J., Imamura, K., Inotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, J., Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Komno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Osato, N., Ota, Y., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toyota, T., Waki, K., Yasunishi, A. and Hayashizaki, Y.

FEATURES
Location/Qualifiers
1..1167
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="J023052021"

ORIGIN

Alignment Scores:
Pred. No.: 4,72e-43 Length: 1167
Score: 691.00 Matches: 157
Percent Similarity: 68.07% Conservative: 37
Best Local Similarity: 55.09% Mismatches: 58
Query Match: 45.28% Indels: 33
DB: 8 Gaps: 5

US-10-660-226-10 (1-305) x AK070318 (1-1167)

QY 1 MetGluAlaGlyValGlyLeuAlaLeuGlnAlaAargAla---AlaGlyPheGlySer 19
Db 179 ATGACGCC---GGTGTGGTCTCCGGCGAAGCGCGCTGGCGCGACTCGGAAC 235
QY 20 SerArgHieArgGlyLeuGlnAlaProThrGlySerLeuArgValAlaAappProAla 39
Db 236 CCACGCCGCTCTTCCACGCCGCGGGTTCGGTG---AGGTTCGCGGTGAGGAAGTTCGCT 292

QY 40 GlyProAlaValAlaValAargAlaAargGlySerLysProValAlaAlaProLeuArgLeuArg 59
Db 233 CAGCGCTTGGTTTGGCGTCCGATCCGCGAGCTCGCGT---GCCAACTGAGGTTTCG 349
QY 60 AlaLysLysSerSerGlyGlyHisGluAenSerHisAenSerValAspGluAlaLeuLeu 79
Db 350 TGTTCACGAGAAACCGCGAGGTATTGACAAACACTTATTACTCTGCTGATGAAGCTCTTGT 409
QY 80 LeuLysArgLysSerGluGluValLeuPheTyLeuAenGlyArgCysIleTyLeuVal 99
Db 410 CTAAGCAAAAAGCAGAGCGTGCCTCTATTGTAATGACCGGTGTATTACTTAGTT 469
QY 100 GlyMetMetGlySerGlyLysSerThrValGlyLysIleMetSerGluValLeuGlyTy 119
Db 470 GGAATGATGGGTCTTGGCAAACTACAGTTGGGAAGATATTAGCTGAAGATTAGGCTAT 529
QY 120 SerPhePheAspSerAspLysLeuValGluGlnAlaValGlyMetProSerValAlaGln 139
Db 530 TCTTCTTTGACAGCGATAAGTTGGTTGAGAAGCGAGTTGGCATATCATCTGTTGCTGAA 589
QY 140 IlePheLysValHisSerGluAlaPhePheArgAspAenGlu-SerSerValLeuArgAs 159
Db 590 ATCTTTTCAGCTCCATAGTGAAGCATTTCTTCAGAGATAATGAGGAGT----- 635
QY 159 pleuSerSerMetArgArgLeuValValAlaThrGlyGlyAlaValIleAtrgProIl 179
Db 635 ----- 635
QY 179 eAsnTrpArgTyrMetLysArgGlyLeuSerValTrpLeuAspValProLeuAspAlaLe 199
Db 636 -----TACATGAAGAAAGGGTCAACAATCTGTTAGATGTTCCATGGATGCACT 685
QY 199 uAlaArgArgIleAlaLysValGlyThrAlaSerArgProLeuLeuAspGlnProSerGl 219
Db 686 TGCAAGACGAATTCGTGCAGTAGGAACGTCTTCGGCCCTCTTTCGCATCAAGAAATCTGG 745
QY 219 yAspProTyrAlaMetAlaPheSerLysLeuSerMetLeuAlaGlnAlaArgGlyAspAl 239
Db 746 TGATCTTATGCAAGGGTTTATGCAAACTGACTGCACCTTTTGGCAAGAAAGATGATC 805
QY 239 aTyAlaAenAlaAspValAargValSerLeuGluLuleAlaCysLysGlnGlyHisAs 259
Db 806 ATATGCTAATGCTGATGCGAGGGTTTCACTCGAACATATTGCAAGTGAACAAGGACATAG 865
QY 259 pAspValSerLysLeuThrProThrAspIleAlaIleGluSerLeuHisLysIleGluSe 279
Db 866 TAATGTGACGACACTTACCAAGTGCCATTGCTATTGAGGCAATTGCTAAAGATGGAG 925
QY 279 rPheValIleGlu 283
Db 926 CTTTCTTACTGAG 938

RESULT 11
AK118899
LOCUS
DEFINITION
Arabidopsis thaliana At2g21940 mRNA for putative shikimate kinase precursor, complete cds, clone: RAFL21-24-A05.
ACCESSION
AK118899
VERSION
AK118899.2 GI:39104604
KEYWORDS
FLI CDNA; CAP trapper.
SOURCE
Arabidopsis thaliana (thale cress)
ORGANISM
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.

REFERENCE
AUTHORS
Seki, M., Iida, K., Satou, M., Sakurai, T., Akiyama, K., Ishida, J., Nakajima, M., Enju, A., Kamiya, A., Narusaka, M., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.
Arabidopsis thaliana full-length cDNA
Published Only in Database (2002)
REFERENCE
2 (bases 1 to 1300)
Seki, M., Iida, K., Satou, M., Sakurai, T., Akiyama, K., Ishida, J.,

KEYWORDS SOURCE ORGANISM

Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 903)

REFERENCE AUTHORS

Kim, C.J., Chen, H., Cheuk, R., Shinn, P., Bowser, L., Carninci, P.,
Chan, M.M., Chang, C.H., Dale, J.M., Hayashizaki, Y., Hsuan, V.W.,
Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J.,
Lam, B., Lee, J.M., Lin, J., Miranda, M., Narusaka, M., Nguyen, M.,
Onodera, C.S., Palm, C.J., Quach, H.L., Sakurai, T., Satou, M., Seki, M.,
Southwick, A., Tang, C.C., Toriumi, M., Wong, C., Wu, H.C., Yamada, K.,
Yu, G., Yuan, S., Shinozaki, K., Davis, R.W., Theologis, A. and
Ecker, J.R.

TITLE JOURNAL

Arabidopsis ORF clones
Unpublished

REFERENCE AUTHORS

2 (bases 1 to 903)
Kim, C.J., Chen, H., Cheuk, R., Shinn, P., Bowser, L., Carninci, P.,
Chan, M.M., Chang, C.H., Dale, J.M., Hayashizaki, Y., Hsuan, V.W.,
Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J.,
Lam, B., Lee, J.M., Lin, J., Miranda, M., Narusaka, M., Nguyen, M.,
Onodera, C.S., Palm, C.J., Quach, H.L., Sakurai, T., Satou, M., Seki, M.,
Southwick, A., Tang, C.C., Toriumi, M., Wong, C., Wu, H.C., Yamada, K.,
Yu, G., Yuan, S., Shinozaki, K., Davis, R.W., Theologis, A. and
Ecker, J.R.

TITLE JOURNAL

Submitted (14-MAR-2003) Salk Institute Genomic Analysis Laboratory
(SIGNAL), Plant Biology Laboratory, The Salk Institute for
Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037,
USA

COMMENT

RIKEN Genomic Sciences Center (GSC) members carried out the
collection and clustering of RAFL cDNAs (RAFL cDNA: 'RIKEN
Arabidopsis Full-length cDNA') : Seki, M., Narusaka, M., Ishida, J.,
Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J.,
Hayashizaki, Y. and Shinozaki, K.

The Salk, Stanford, PGEN (SSP) Consortium members constructed and
sequenced the pUNI (ORF) clones using the RAFL cDNAs: Kim, C.J.,
Chen, H., Cheuk, R., Shinn, P., Bowser, L., Chan, M.M., Chang, C.M.,
Dale, J.M., Hsuan, V.W., Jones, T., Karlin-Neumann, G., Lam, B.,
Lee, J.M., Lin, J., Miranda, M., Nguyen, M., Onodera, C.S., Palm, C.J.,
Quach, H.L., Southwick, A., Tang, C.C., Toriumi, M., Wong, C., Wu, H.C.,
Yamada, K., Yu, G., Yuan, S., Davis, R.W., Theologis, A., and
Ecker, J.R.

Kim, C.J. (SSP/Salk) and Seki, M. (RIKEN GSC) contributed equally to
this work. Shinozaki, K. (RIKEN GSC) and Ecker, J.R. (SSP/Salk)
contributed equally to this work as PIs.

FEATURES source

Location/Qualifiers
1..903
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosomes="4"
/clone="U60448"
/ecotype="Columbia"
/notes="This clone is in pUNI 51"

CDS

1..903
/notes="putative shikimate kinase"
/codon_start=1
/product="At4g39540"
/protein_id="AA063355.1"
/db_xref="GI:28950863"

/translation="MEATVORPOYSSWDLNFEKPKRSLYNTORIKEDKFRVV
ALTLDKRHRRLRSVDKSSALLETGSLHSPFDEQIILKKAEVYKPYLNGSMY
LVGMGSGKTTVGKINWARSIGVTFDFCDTLIIOAMKGTVSABIFHFSGVREKETE
ALKKLSIMYHVVGGGAVIRPINWYMKHGISIWLDVPLEAHLAAVGTGSRP
LLHDDSGGTYTAAALNRLSTIWDARGEATYKASARVSLNITLKLGRVSDLTDAEI
AIEAFQVQSYLEKEDGMARPDGL"

ORIGIN

Alignment Scores:

Pred. No.:	8,278-37	Length:	903
Score:	606.50	Matches:	130
Percent Similarity:	71.37%	Conservative:	37
Best Local Similarity:	55.56%	Mismatches:	58
Query Match:	39.74%	Indels:	9
DB:	8	Gaps:	5
US-10-660-226-10 (1-305) x BT005291 (1-903)			
QY	57	ArgLeuArgAla-----LysLysSerSerGlyGlyHisGluAnSer-----	70
Db	163	CGACTTAGATCAGTTTCTGATAGAACTCTCAGCATTGTTGGAACTGGAAGTCTTCTT	222
QY	71	HisAsnSerValasp---GluAlaLeuLeuLysArgLysSerGluGluValLeuPhe	89
Db	223	CATTCTCCATTGATGAAGAAACACAGATTGTAAGAAAAAGCTGAAGAGGTAAACCG	282
QY	90	TyrLeuAnGlyValGlyCysIleTyrLeuValGlyMetMetGlySerGlySerThrVal	109
Db	283	TATTTAAATGACGATCGATGATCTTGTGTATGATGGGTCCGGGAAACGACTGTA	342
QY	110	GlyLysIleMetSerGluValLeuGlyTyrSerPhePheAspSerAspLysLeuValGlu	129
Db	343	GGGAGATTATGGCAAGATCGCTTGTTATACATCTTTGATTGTGACACTTTGATCGAG	402
QY	130	GlnAlaValGlyMetProSerValAlaGlnIlePheLysValHisSerGluAlaPhePhe	149
Db	403	CAGGCTATGAGGAACTCTCTGTAGCTGAGATATTGAGCATTTCCGTGAGAGTGTCTTC	462
QY	150	ArgAspAsnGluSerSerValLeuArgAspLeuSerSerMet---ArgArgLeuValVal	168
Db	463	AGAGAAAAGAGACTGAAGCGTTAAAGAACTCTCTTGTATGATACCAACCAAGTTGTGT	522
QY	169	AlaThrGlyGlyGlyAlaValIleArgProIleAsnTyrArgTyrMetLysArgGlyLeu	188
Db	523	TCAACCGGGGCGGCGAGTTATAGACCCATCAATTGGAAGTACATGATTAAGGTATT	582
QY	189	SerValTyrLeuAspValProLeuAspAlaLeuAlaArgIleAlaLysValGlyThr	208
Db	583	AGTATTTCGCTTGTAGTACCTCTAGAAGCTTTAGGCGCATAGATAAGTCTCTAGAACT	642
QY	209	AlaSerArgProLeuLeu---AspGlnProSerGlyAspProTyrAlaMetAlaPheSer	227
Db	643	GTTTCAAGACCATCTCTCATGATGATGAGTGGGGGACACATACACAGCGGCTTTAAAC	702
QY	228	LysLeuSerMetLeuAlaGlnGlnArgGlyAspAlaTyrAlaAsnAlaAspValArgVal	247
Db	703	CGTCTTTCAACGATTGGATGCCAGTGGTGAAGCATACACTAAAGCCAGCGCAAGATT	762
QY	248	SerLeuGluGluIleAlaCysLysGlnGlyHisAspAspValSerLysLeuThrProThr	267
Db	763	TCCTTGGAGATATTACTTTGAAGCTCGTTATAGAAGTGTCTCAGATCTTACACAGCT	822
QY	268	AspIleAlaIleGluSerLeuHisLysIleGluSerPheVal	281
Db	823	GAATCGCCATTGAGGCTTTGAGCAAGTTCAGAGCTATTTA	864
RESULT 15			
AK117791			
LOCUS			
DEFINITION			
Arabidopsis thaliana At4g39540 mRNA for putative shikimate kinase,			
complete cds, clone: RAFL17-48-C24.			
AK117791			
ACCESSION			
VERSION			
AK117791.1 GI:26450650			
KEYWORDS			
FLI CDNA; CAP trapper.			
SOURCE			
ORGANISM			
Arabidopsis thaliana (thale cress)			
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.			
REFERENCE			
AUTHORS			
Seki, M., Tida, K., Satou, M., Sakurai, T., Akiyama, K., Ishida, J., Nakajima, M., Enju, A., Kamiya, A., Narusaka, M., Carninci, P.,			

Kawai,J., Hayashizaki, Y. and Shinozaki, K.
 Arabidopsis thaliana full-length cDNA
 Published Only in Databse (2002)
 2 (bases 1 to 1294)
 Seki,M., Iida,K., Satou,M., Sakurai,T., Akiyama,K., Ishida,J.,
 Nakajima,M., Enju,A., Kamiya,A., Narusaka,M., Carninci,P.,
 Kawai,J., Hayashizaki, Y. and Shinozaki, K.
 Direct Submission
 Submitted (25-NOV-2002) Motoaki Seki, RIKEN Genomic Sciences
 Center; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa
 230-0045, Japan [E-mail: mseki@gsc.riken.go.jp,
 URL: http://pfjweb.gsc.riken.go.jp, Tel: 81-45-503-9625,
 Fax: 81-45-503-9586]
 An Arabidopsis full-length cDNA library was constructed essentially
 as reported previously (Seki et al. (1998) Plant J. 15:707-720;
 Seki et al. (2002) Science 296:141-145). cDNA cleaved with BamHI
 and XhoI was ligated to modified Lambda FLC-1-E vector (Carninci et
 al. (2001) Genomics 77:79-90) digested with BamHI and SalI.
 This clone is in a modified pBluescript vector.
 Please visit our web site (http://pfjweb.gsc.riken.go.jp/) for
 further details.

FEATURES

Location/Qualifiers

1..1294
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /db_xref="taxon:3702"
 /chromosomes="4"
 /clone="RAF17-48-C24"
 /ecotype="Columbia"
 /note="common name: thale cress"

gene

1..1294
 /gene="At4g39540/F23K16_170"

CDS

228..1130
 /gene="At4g39540/F23K16_170"
 /codon_start=1
 /product="putative shikimate kinase"
 /protein_id="BAC42436.1"
 /db_xref="GI:26450651"
 /translation="MEATVQRFQYSSWNLNPFQKPRGSLRYNTQRIKEDKRRFVV
 ALTKDRRLRSVSDNSALLETSLLHSFPDEBQQILKKAEVVKPYLNGRSMY
 LVGMGSGTGVGKIMARSLGYTFDCDTLI EQAMKGTSAEIPFHFGESVFEKETE
 ALKLSLVHYVGVGGAVRPINWKYMHKIGISILWLDVPLELAHRIAAGVGTGRP
 LHHDESGDTYTAALNRLSTIWDARGEAYTKASARVLENITLKLGRSVSDLTPAEI
 ATEAPEQVQSYLEKEDGMARPDGL"

ORIGIN

Alignment Scores:
 Pred. No.: 1.26e-36 Length: 1294
 Score: 606.50 Matches: 130
 Percent Similarity: 71.37% Conservative: 37
 Best Local Similarity: 55.56% Mismatches: 58
 Query Match: 39.74% Indels: 9
 DB: 8 Gaps: 5

US-10-660-226-10 (1-305) x AK117791 (1-1294)

QY 57 ArgLeuArgAla-----LysLysSerSerGlyGlyHisGluAsnSer----- 70
 DB 390 CGACTTAGATCATGTTCTCTGATAAGAACTCTCAGCATTTGTTGGAAACTCGAAGTCTTCTT 449
 QY 71 HisAsnSerValAsp--GluAlaLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuPhe 89
 DB 450 CATTCCTCCATTGATGAAGACACACAGATTTGAGAGAAAAGCTGAAGAGGTAAACCG 509
 QY 90 TyrLeuAsnGlyArgCysIleTyrLeuValGlyMetMetGlySerGlyLysSerThrVal 109
 DB 510 TATTTAAATGGACATCGATGATCTTTGTTGATGATGGTTCGGGAAACGACTGTA 569
 QY 110 GlyLysIleMetSerGluValLeuGlyTyrSerPhePheAspSerAspLysLeuValGlu 129
 DB 570 GGGAAGATTATGGCAAGATCGCTTGTGTTATACATCTTTGATGTGACACTTTGATCGAG 629
 QY 130 GlnAlaValGlyMetProSerValAlaGlnIlePheLysValHisSerGluAlaPhePhe 149

Db 630 CAGGCTATGAAGGAACTTCTGTAGCTAGATATTTGAGCATTTTCGGTGAGAGTCTCTTC 689
 QY 150 ArgAspAenGluSerSerValLeuArgAspLeuSerSerMet---ArgArgLeuValVal 168
 Db 690 AGAGAAAAAGAGACTGAAGCGTTAAAGAACTCTCTTTGATGTACCAACCAAGTTGTGTT 749
 QY 169 AlaThrGlyGlyAlaValIleArgProIleAsnTTPArgTyrMetLysArgGlyLeu 188
 Db 750 TCAACCGGGGGGGGGCGAGTTATAGACCCATCAATTGGAAGTACATCATATAAAGTATT 809
 QY 189 SerValTTPLeuAspValProLeuAspAlaLeuAlaArgArgIleAlaLysValGlyThr 208
 Db 810 AGTATTTCGTTGATGTAACCTCTAGAACCTTAGCGCATAGATAGCTGTAGGAAC 869
 QY 209 AlaSerArgProLeuLeu---AspGlnProSerGlyAspProTyrAlaMetAlaPheSer 227
 Db 870 GGTTCAGACCACTTGCTACATGATGATGATCAGGCGGACACATACACAGCGCTTTAAAC 929
 QY 228 LysLeuSerMetLeuAlaGlnGlnArgGlyAspAlaTyrAlaAsnAlaAspValArgVal 247
 Db 930 CGTCTTTCAACGATTTGGGATGCAGTGTGGAAGCATACACTAAAGCCAGCGCAAGATT 989
 QY 248 SerLeuGluGluIleAlaCysLysGlnGlyHisAspAspValSerLysLeuThrProThr 267
 Db 990 TCCITGGAGAAATATTACTTTGAAGCTCGGTTATAGAAAGTGCTCAGATCTTACACCAGCT 1049
 QY 268 AspIleAlaIleGluSerLeuHisLysIleGluSerPheVal 281
 Db 1050 GAAATCGCCATTGAGGCGCTTTGAGCAGAGTTTCAGAGCTATTATA 1091

Search completed: August 25, 2005, 22:17:38

Job time : 5030 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 25, 2005, 17:38:27 ; Search time 610 Seconds

(without alignments)
2959,871 Million cell updates/sec

Title: US-10-660-226-10

Perfect score: 1526

Sequence: 1 MEAGGVLQALARAAGFGSS.....ADSSADQAQESQIRIQTL 305

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DRV=xlh
-Q=/cgn2_1/USPTO_spool/US10660226/runat_25082005_110426_28514/app_query.fasta_1.455
-DB=N_Geneseq_16Dec04 -QFIX=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blom62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFWT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10660226@cgn 1.1 470 @runat_25082005_110426_28514 -NCPU=6 -ICPU=3
-NO_WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq_16Dec04:*

1:	Geneseqn1980s:*
2:	Geneseqn1990s:*
3:	Geneseqn2000s:*
4:	Geneseqn2001as:*
5:	Geneseqn2001bs:*
6:	Geneseqn2002as:*
7:	Geneseqn2002bs:*
8:	Geneseqn2003as:*
9:	Geneseqn2003bs:*
10:	Geneseqn2003cs:*
11:	Geneseqn2003ds:*
12:	Geneseqn2004as:*
13:	Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	660	43.3	912	3 AAC50023	Aac50023 Arabidops
2	656	43.0	1363	3 AAC36026	Aac36026 Arabidops
3	650	42.6	1365	3 AAC50026	Aac50026 Arabidops
4	650	42.6	1491	3 AAC50029	Aac50029 Arabidops
5	631.5	41.4	637	10 ADK52674	Adk52674 Plant DNA

6	626.5	41.1	1484	3 AAC40184	Aac40184 Arabidops
7	602	39.4	1095	3 AAC47100	Aac47100 Arabidops
8	576.5	37.8	1093	3 AAC36746	Aac36746 Arabidops
9	530.5	34.8	696	3 AAC41622	Aac41622 Arabidops
10	473	31.0	1010	13 ADR60835	Adr60835 Cotton CD
11	385.5	25.3	660	10 ADK58257	Adk58257 Plant DNA
12	374	24.5	309	6 ABN79354	Abn79354 Human kin
13	368	24.1	329	10 ADK58255	Adk58255 Plant DNA
14	362.5	23.8	614	13 ADR60836	Adr60836 Cotton CD
15	347.5	22.8	573	13 ACN52580	Acn52580 Cotton an
16	318.5	20.9	570	13 ADS47877	Ads47877 Bacterial
17	315	20.6	564	13 ADT45202	Adt45202 Bacterial
18	307	20.1	528	13 ADT46243	Adt46243 Bacterial
19	294.5	19.3	402	13 ADT44359	Adt44359 Bacterial
20	271.5	17.8	537	10 ADK58256	Adk58256 Plant DNA
21	256	16.8	226	7 ADK67657	Adk67657 Corn seed
22	254	16.6	5330	2 AAX84337	Aax84337 Stealth v
23	252	16.5	582	13 ADT45821	Adt45821 Bacterial
24	249.5	16.3	507	13 ADK5422	Adk5422 Bacterial
25	245	16.1	696	12 ADL02481	Adl02481 DNA encod
26	245	16.1	14335	4 AAF28524	Aaf28524 Genomic f
27	241	15.8	678	12 ADM34080	Adm34080 Cryptospo
28	237.5	15.6	1479	13 ADS48191	Ads48191 Bacterial
29	234.5	15.4	603	13 ADT42262	Adt42262 Bacterial
30	231	15.1	543	13 ADK31460	Adk31460 Shikimate
31	228.5	15.0	504	13 ADS56012	Ads56012 Bacterial
32	227.5	14.9	495	13 ADS56789	Ads56789 Bacterial
33	226.5	14.8	480	10 ADE82174	Ade82174 Arabidops
34	222.5	14.6	498	13 ADS57732	Ads57732 Bacterial
35	222	14.5	603	9 ADA30736	Ada30736 DNA encod
36	222	14.5	603	11 ACH97846	Ach97846 Klebsiell
37	220	14.4	438	13 ADT46957	Adt46957 Bacterial
38	220	14.4	444	13 ADT43049	Adt43049 Bacterial
39	218	14.3	513	10 ABZ38655	Abz38655 N. gonorr
40	218	14.3	56485	3 AA881476	Aa881476 N. mening
41	218	14.3	110000	3 AA881489_4	Continuation (5 of
42	218	14.3	349980	3 AAF21612	Aaf21612 Neisseria
43	217.5	14.3	577	13 ACN52511	Acn52511 Cotton an
44	217	14.2	522	10 ACF70327	Acf70327 Phototrab
45	217	14.2	110000	10 ACF65385_5	Continuation (6 of

ALIGNMENTS

RESULT 1
AAC50023
ID AAC50023 standard; DNA; 912 BP.
XX
AC AAC50023;
XX
18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 63305.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic pathway;
KW promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.

PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0133256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140595P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142202P.
PR 12-JUL-1999; 99US-0142377P.
PR 13-JUL-1999; 99US-0143342P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.

PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145227P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 04-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149923P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158023P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.

PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142190P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145324P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145513P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.

PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152263P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Alignment Scores:

Pred. No.:	1,11e-53	Length:	1363
Score:	656.00	Matches:	127
Percent Similarity:	79.80%	Conservative:	35
Best Local Similarity:	62.56%	Mismatches:	41
Query Match:	42.99%	Indels:	0
DB:	3	Gaps:	0

```
US-10-660-226-10 (1-305) x AAC36026 (1-1363)
QY 79 LeuLeuLysArgLysSerGluGluValLeuPheTyrLeuAsnGlyArgCysIleTyrLeu 98
Db 368 ATTCTCAAGAGAAAGCAGAGAGGTTAAACCGTATTTGATGGACGATCTATGTACCTT 427
QY 99 ValGlyMetMetGlySerGlyLysSerThrValGlyLysIleMetSerGluValLeuGly 118
Db 428 GTCGAATGATGGTTCTCGGAAACCACTGTGGGAAAGTTAATGTCCAAAGTGTCTCGGT 487
QY 119 TyrSerPhePheAspSerAspLysLeuValGluAlaValGlyMetProSerValAla 138
Db 488 TATACGTCTCTTCACATCGCAGCACTTTGATTGAACAGCGCATGAATGGAACCTTCTGTGCA 547
QY 139 GluIlePheLysValHisSerGluAlaPhePheArgAspAsnGluSerSerValLeuArg 158
Db 548 GAGATATTTGTTTCATACGGAGAGAAATTTTTAGAGGAAGAGACCGATCGCCTTAAG 607
QY 159 AspLeuSerSerMetArgArgLeuValValAlaThrGlyGlyAlaValIleArgPro 178
Db 608 AAGCTCTCTTCGAGGTATCAAGTTGTTGTTCCACAGGTGGAGTGCAGTTATAAGACCC 667
QY 179 IleAsnTrpArgTyrMetLysArgGlyLeuSerValTrpLeuAspValProLeuAspAla 198
Db 668 ATTAAGTGAAGATATATGCATAAAGGAATCAGCAATTTGGCTAGATGTCCTCTAGAAGCA 727
QY 199 LeuAlaArgArgIleAlaLysValGlyThrAlaSerArgProLeuLeuAspGlnProSer 218
Db 728 TTAGCCCATAGATCGCTGCTGTGTGAACATGATTCAGACCACTGCTACAGATGATCA 787
QY 219 GlyAspProTyrAlaMetAlaPheSerLysLeuSerMetLeuAlaGlnArgGlyAsp 238
Db 788 GGAGATGCTACTACTCAGTGGCTTTCAAACGCTCTCTCGCTATTGGGACGAGCGGTGAA 847
QY 239 AlaTyrAlaAsnAlaAspValArgValSerLeuGluGluIleAlaCysLysGlnGlyHis 258
Db 848 GCATACACAAACGCAATGCCAGAGTCTCTTAGAAAAATATTGCAGCAAGCGTGGCTAT 907
QY 259 AspAspValSerLysLeuThrProThrAspIleAlaIleGluSerLeuHisLysIleGlu 278
Db 908 AAAAATGCTCAGATCTCACACCACTGAAATTTGTATCGAGCCCTTCGAGCAAGTTCG 967
QY 279 SerPheVal 281
Db 968 AGCTTCTA 976

RESULT 3
AAC50026
ID AAC50026 standard; DNA; 1365 BP.
XX
AC AAC50026;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 63316.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic pathway;
KW promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
```

```
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0132866P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 02-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142300P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
```

PR 16-JUL-1999;	99US-0144086P.	PR 08-OCT-1999;	99US-0158232P.
PR 19-JUL-1999;	99US-0144325P.	PR 12-OCT-1999;	99US-0158369P.
PR 19-JUL-1999;	99US-0144331P.	PR 13-OCT-1999;	99US-0158293P.
PR 19-JUL-1999;	99US-0144332P.	PR 13-OCT-1999;	99US-0159294P.
PR 19-JUL-1999;	99US-0144333P.	PR 13-OCT-1999;	99US-0159295P.
PR 19-JUL-1999;	99US-0144334P.	PR 14-OCT-1999;	99US-0159329P.
PR 19-JUL-1999;	99US-0144335P.	PR 14-OCT-1999;	99US-0159330P.
PR 20-JUL-1999;	99US-0144352P.	PR 14-OCT-1999;	99US-0159331P.
PR 20-JUL-1999;	99US-0144632P.	PR 14-OCT-1999;	99US-0159637P.
PR 20-JUL-1999;	99US-0144884P.	PR 14-OCT-1999;	99US-0159638P.
PR 21-JUL-1999;	99US-0144814P.	PR 18-OCT-1999;	99US-0159584P.
PR 21-JUL-1999;	99US-0145086P.	PR 21-OCT-1999;	99US-0160741P.
PR 21-JUL-1999;	99US-0145088P.	PR 21-OCT-1999;	99US-0160767P.
PR 22-JUL-1999;	99US-0145085P.	PR 21-OCT-1999;	99US-0160768P.
PR 22-JUL-1999;	99US-0145087P.	PR 21-OCT-1999;	99US-0160770P.
PR 22-JUL-1999;	99US-0145089P.	PR 21-OCT-1999;	99US-0160814P.
PR 22-JUL-1999;	99US-0145192P.	PR 21-OCT-1999;	99US-0160815P.
PR 23-JUL-1999;	99US-0145145P.	PR 22-OCT-1999;	99US-0160980P.
PR 23-JUL-1999;	99US-0145218P.	PR 22-OCT-1999;	99US-0160981P.
PR 23-JUL-1999;	99US-0145224P.	PR 22-OCT-1999;	99US-0160989P.
PR 26-JUL-1999;	99US-0145276P.	PR 25-OCT-1999;	99US-0161404P.
PR 27-JUL-1999;	99US-0145913P.	PR 25-OCT-1999;	99US-0161403P.
PR 27-JUL-1999;	99US-0145918P.	PR 25-OCT-1999;	99US-0161406P.
PR 27-JUL-1999;	99US-0145919P.	PR 26-OCT-1999;	99US-0161359P.
PR 28-JUL-1999;	99US-0145951P.	PR 26-OCT-1999;	99US-0161360P.
PR 02-AUG-1999;	99US-0146386P.	PR 26-OCT-1999;	99US-0161361P.
PR 02-AUG-1999;	99US-0146388P.	PR 28-OCT-1999;	99US-0161920P.
PR 02-AUG-1999;	99US-0146389P.	PR 28-OCT-1999;	99US-0161992P.
PR 03-AUG-1999;	99US-0147038P.	PR 28-OCT-1999;	99US-0161993P.
PR 04-AUG-1999;	99US-0147204P.	PR 28-OCT-1999;	99US-0161993P.
PR 04-AUG-1999;	99US-0147302P.	PR 29-OCT-1999;	99US-0162142P.
PR 05-AUG-1999;	99US-0147192P.	Alignment Scores:	
PR 05-AUG-1999;	99US-0147260P.	Pred. No.:	4, 228-53
PR 06-AUG-1999;	99US-0147303P.	Score:	650.00
PR 06-AUG-1999;	99US-0147416P.	Percent Similarity:	80.10%
PR 09-AUG-1999;	99US-0147493P.	Best Local Similarity:	62.69%
PR 09-AUG-1999;	99US-0147935P.	Query Match:	42.60%
PR 10-AUG-1999;	99US-0148171P.	DB:	3
PR 11-AUG-1999;	99US-0148319P.	US-10-660-226-10 (1-305) x AAC50026 (1-1365)	
PR 12-AUG-1999;	99US-0148341P.	QY 81	LysArgLysSerGluGluValLeuPheTyrLeuAsnGlyArgCysaileTyrLeuValGly 100
PR 13-AUG-1999;	99US-0148565P.	Db 378	CAAGAAAGCAGAGAGGTTAAACCGTATTTGAATGGACGATCTATGTACCTTGTTCGGA 437
PR 16-AUG-1999;	99US-0149368P.	QY 101	MetMetGlySerGlyLysSerThrValGlyLysIleMetSerGluValLeuGlyTyrSer 120
PR 17-AUG-1999;	99US-0149175P.	Db 438	ATGATGGGTTCTGGGAAAAACAACCTGGGAAAGTTAATGTCCAAAAGTGTCTCGTTATACG 497
PR 18-AUG-1999;	99US-0149426P.	QY 121	PhePheAspSerAspLysLeuValGluGlnAlaValGlyMetProSerValAlaGlnIle 140
PR 20-AUG-1999;	99US-0149722P.	Db 498	TTCTTTGACTGGCACACTTTGATTTGAACAGGCGATGAATGGAACTCTCTTTCAGAGATA 557
PR 20-AUG-1999;	99US-0149723P.	QY 141	PheLysValHisSerGluAlaPhePheArgAspAsnGluSerSerValLeuArgAspLeu 160
PR 23-AUG-1999;	99US-0149902P.	Db 558	TTTGTTCATCAGGAGAGAAATTTTTTAGAGGAAAGAGACCGCGCTTAAGAAGCTC 617
PR 25-AUG-1999;	99US-0150566P.	QY 161	SerSerMetArgArgLeuValValAlaThrGlyGlyAlaValIleArgProIleAsn 180
PR 26-AUG-1999;	99US-0150884P.	Db 618	TCTTCGAGGTATCAAGTTGTTGTTCCACAGGTGGAGGTGCAGTTATATAAGACCATTAAC 677
PR 27-AUG-1999;	99US-0151065P.	QY 181	TtpArgTyrMetLysArgGlyLeuSerValTtpLeuAspValProLeuAspAlaLeuAla 200
PR 27-AUG-1999;	99US-0151066P.	Db 678	TGGAGTATATGCATAAAGGAATCAGCATTTGGTAGATGTGCTCTAGAACATTAGCC 737
PR 27-AUG-1999;	99US-0151080P.	QY 201	ArgArgIleAlaLysValGlyThrAlaSerArgProLeuLeuAspGlnProSerGlyAsp 220
PR 30-AUG-1999;	99US-0151303P.	Db 738	CATAGAATCGCTGCTGTGGAACTGATTCACGCCACTGCTACACGATGAATCAGAGAT 797
PR 31-AUG-1999;	99US-0151438P.	QY 221	ProTyrAlaMetAlaPheSerLysLeuSerMetLeuAlaGlnArgGlyAspAlaTyr 240
PR 01-SEP-1999;	99US-0151930P.	Db 798	GCATACTCAGTGGCTTCAACAGCTCTCTCGGTATTTTGGGACGAGCGGTGAAGCATAC 857
PR 10-SEP-1999;	99US-0152363P.		
PR 10-SEP-1999;	99US-0153070P.		
PR 13-SEP-1999;	99US-0153758P.		
PR 15-SEP-1999;	99US-0154018P.		
PR 16-SEP-1999;	99US-0154039P.		
PR 20-SEP-1999;	99US-0154779P.		
PR 22-SEP-1999;	99US-0155139P.		
PR 23-SEP-1999;	99US-0155486P.		
PR 24-SEP-1999;	99US-0155659P.		
PR 28-SEP-1999;	99US-0156458P.		
PR 29-SEP-1999;	99US-0156596P.		
PR 04-OCT-1999;	99US-0157117P.		
PR 05-OCT-1999;	99US-0157753P.		
PR 06-OCT-1999;	99US-0157865P.		
PR 07-OCT-1999;	99US-0158029P.		

Qy 241 AlaAsnAlaAspValIArgValSerLeuGluGluIleAlaCysLysGlnGlyHisAspAsp 260
Db 858 ACAACGCAAAATGCCAGGTCCTTAGAAAAATATTGCAGCAAGCGCTGCTATAAAAAAT 917
Qy 261 ValSerLysLeuThrProThrAspIleAlaIleGluSerLeuHisLysIleGluSerPhe 280
Db 918 GTCTCAGATCTACACCAACTGAATTCGTATCGAGGCCCTTCGACGAAGTTCTGAGCTTT 977
Qy 281 Val 281
Db 978 CTA 980
RESULT 4
AAC50029
ID AAC50029 standard; DNA; 1491 BP.
XX
AC AAC50029;
XX
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 63328.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic pathway;
KW promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0122380P.
PR 09-MAR-1999; 99US-0123348P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140153P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144684P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.

PD 13-MAR-2003.
 XX 30-AUG-2002; 2002WO-US027884.
 XX 31-AUG-2001; 2001US-0316471P.
 XX (DOWC) DOW CHEM CO.
 PA (DOWC) DOW AGROSCIENCES LLC.
 XX
 PI Weglarz T, Gachotte D, Blakeslee B, McCreary DA, Pell RJ;
 PI Oriedo JVB, Crosley R, Reddy AS, Shukla V, Larrinua I, Miller BA;
 XX WPI; 2003-313091/30.
 XX
 XX Novel genes that confer altered metabolic characteristics in Nicotiana
 PT benthamiana plants, useful for altering the levels of metabolites e.g.
 PT acids, fatty acids, amino acids, carbohydrates, hydrocarbons and sterols.
 XX
 PS Claim 1; SEQ ID NO 57; 2576pp; English.
 XX
 CC The invention comprises DNA sequences which confer an altered metabolic
 CC characteristic when they are expressed in a plant. The DNA sequences of
 CC the invention are useful for producing plants with an altered metabolic
 CC characteristic, such as: altered acid metabolism, alcohol metabolism,
 CC fatty acid metabolism, branched fatty acid metabolism, alkaloid or other
 CC base metabolism, altered amino acid metabolism, altered ester metabolism,
 CC altered glyceride metabolism, altered phenolic metabolism, altered
 CC carbohydrate metabolism, altered sterol, oxygenated terpene, or
 CC isoprenoid metabolism, alkene or alkene metabolism, hydrocarbon
 CC metabolism, ketone or quinone metabolism. The DNA sequences of the
 CC invention may be used to provide disease resistance in a plant and gene
 CC shuffling or sexual PCR procedures. The present nucleic acid represents a
 CC DNA sequence of the invention.
 XX
 SQ Sequence 637 BP; 185 A; 103 C; 165 G; 184 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 9,42e-52 Length: 637
 Score: 631.50 Matches: 121
 Percent Similarity: 80.73% Conservative: 34
 Best Local Similarity: 63.02% Mismatches: 36
 Query Match: 41.38% Indels: 1
 DB: 10 Gaps: 1

US-10-660-226-10 (1-305) x ADK52674 (1-637)

QY 90 TyrLeuAsnGlyArgCysIleTyrLeuValGlyMetMetClySerGlyLysSerThrVal 109
 Db 9 TATCTAGATGGACGATGTATATACCTCGTTGGAGTATCGCTCGCAAAAACACTGTG 68
 QY 110 GlyLysIleMetSerGluValLeuGlyTyrSerPheAspSerAspLysLeuValGlu 129
 Db 69 GCGCGTATTGGCAGAAACACTGGGATATCTCTTTTGACTGTGACAGGCTGATAGAG 128
 QY 130 GlnAlaValGlyMetProSerValAlaGlnIlePheLysValHisSerGluAlaPhePhe 149
 Db 129 CAGGCTGTTGGTGGAACTACAGTAGCTGAATCTTCAAGCTTCGTGGAGAGAGCTTCTTT 188
 QY 150 ArgAspAsnGluSerSerValLeuArgAspLeuSerSerMetArgLeuValAlaAla 169
 Db 189 AGGGACAATGAGACGAGGAGGTATTGCAACAGCTGTCTGCGATGATCGGCTTGTGTTCA 248
 QY 170 ThrGlyGlyValAlaValIleArgProIleAsnTyrArgTyrMetLysArgGlyLeuSer 189
 Db 249 ACAGGTGGAGGTGCAGTTGTTCTCCATTAAATTGGACATATGCAACAGGATATTAGT 308
 QY 190 ValTrpLeuAspValProLeuAspAlaLeuAlaArgArgIleAlaLysValGlyThrAla 209
 Db 309 GTTTGGTGTAGATCTCTCTTAGATGCTTTGGCCAAGAGAGTTACTCTCTGAAGGAATAA 368
 QY 210 SerArgProLeuLeuAspGlnProSerGlyAspProTyrAlaMetAlaPheSerLysLeu 229
 Db 369 TCTCGACCCCTATTACATGAAGAATCAGGAGACATTTATGATAAGACTTTGAAGCGGTTA 428

QY 230 SerMetLeuAlaGlnGlnArgGlyAspAlaTyrAlaAsnAlaAspValArgValSerLeu 249
 Db 429 ACTACTTTTAATGGAGACAAGGGGTGAAAACACTATGCCAATGCAAGTGTCAAGGTTTCACCTA 488
 QY 250 GluGluIleAlaCysLysGlnGlyHisAspValSerLysLeuThrProThrAspIle 269
 Db 489 GAAATATATTGCAGTGAAGAGGGA---AAAGATGTCTGCCATATTATACACCTACTGAAT 545
 QY 270 AlaIleGluSerLeuHisLysIleGluSerPheVal 281
 Db 546 ACTTAGAGTCTTATATCAAAATTCAGAACTTCTTA 581
 RESULT 6
 AAC40184
 ID AAC40184 standard; DNA; 1484 BP.
 XX AAC40184;
 XX
 DT 17-OCT-2000 (first entry)
 XX
 DE Arabidopsis thaliana DNA fragment SEQ ID NO: 27368.
 XX
 KW Hybridisation assay; genetic mapping; gene expression control;
 KW protein identification; signal transduction pathway; metabolic pathway;
 KW promoter; termination sequence; ss.
 XX
 OS Arabidopsis thaliana.
 XX
 FN EP1033405-A2.
 XX
 PD 06-SEP-2000.
 XX
 PF 25-FEB-2000; 2000EP-00301439.
 XX
 PR 25-FEB-1999; 99US-0121825P.
 PR 05-MAR-1999; 99US-0123180P.
 PR 09-MAR-1999; 99US-0123548P.
 PR 23-MAR-1999; 99US-0125788P.
 PR 25-MAR-1999; 99US-0126264P.
 PR 29-MAR-1999; 99US-0126785P.
 PR 01-APR-1999; 99US-0127462P.
 PR 06-APR-1999; 99US-0128234P.
 PR 08-APR-1999; 99US-0128714P.
 PR 16-APR-1999; 99US-0129845P.
 PR 19-APR-1999; 99US-0130077P.
 PR 21-APR-1999; 99US-0130449P.
 PR 23-APR-1999; 99US-0130510P.
 PR 28-APR-1999; 99US-0130891P.
 PR 30-APR-1999; 99US-0131449P.
 PR 30-APR-1999; 99US-0132048P.
 PR 30-APR-1999; 99US-0132407P.
 PR 04-MAY-1999; 99US-0132484P.
 PR 05-MAY-1999; 99US-0132485P.
 PR 06-MAY-1999; 99US-0132486P.
 PR 06-MAY-1999; 99US-0132487P.
 PR 07-MAY-1999; 99US-0132863P.
 PR 11-MAY-1999; 99US-0134256P.
 PR 14-MAY-1999; 99US-0134218P.
 PR 14-MAY-1999; 99US-0134219P.
 PR 14-MAY-1999; 99US-0134221P.
 PR 14-MAY-1999; 99US-0134370P.
 PR 18-MAY-1999; 99US-0134768P.
 PR 19-MAY-1999; 99US-0134941P.
 PR 20-MAY-1999; 99US-0135124P.
 PR 21-MAY-1999; 99US-0135353P.
 PR 24-MAY-1999; 99US-0135629P.
 PR 25-MAY-1999; 99US-0136021P.
 PR 27-MAY-1999; 99US-0136392P.
 PR 28-MAY-1999; 99US-0136782P.
 PR 01-JUN-1999; 99US-0137222P.
 PR 03-JUN-1999; 99US-0137528P.
 PR 04-JUN-1999; 99US-0137502P.

PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 14-JUN-1999; 99US-0138847P.
PR 16-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 17-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142277P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144086P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.

PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-015753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160880P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Alignment Scores:
Pred. No.:
Score:

8.93e-51
626.50

Length:
Matches:

1484
126

PR	13-JUL-1999;	99US-0143542P.
PR	14-JUL-1999;	99US-0143624P.
PR	15-JUL-1999;	99US-0144005P.
PR	16-JUL-1999;	99US-0144086P.
PR	19-JUL-1999;	99US-0144325P.
PR	19-JUL-1999;	99US-0144331P.
PR	19-JUL-1999;	99US-0144332P.
PR	19-JUL-1999;	99US-0144333P.
PR	19-JUL-1999;	99US-0144334P.
PR	19-JUL-1999;	99US-0144335P.
PR	20-JUL-1999;	99US-0144632P.
PR	20-JUL-1999;	99US-0144884P.
PR	21-JUL-1999;	99US-0144814P.
PR	21-JUL-1999;	99US-0145086P.
PR	21-JUL-1999;	99US-0145088P.
PR	22-JUL-1999;	99US-0145087P.
PR	22-JUL-1999;	99US-0145089P.
PR	22-JUL-1999;	99US-0145192P.
PR	23-JUL-1999;	99US-0145218P.
PR	23-JUL-1999;	99US-0145224P.
PR	26-JUL-1999;	99US-0145276P.
PR	27-JUL-1999;	99US-0145813P.
PR	27-JUL-1999;	99US-0145918P.
PR	27-JUL-1999;	99US-0145919P.
PR	28-JUL-1999;	99US-0145951P.
PR	02-AUG-1999;	99US-0146386P.
PR	02-AUG-1999;	99US-0146388P.
PR	03-AUG-1999;	99US-0146389P.
PR	04-AUG-1999;	99US-0147204P.
PR	05-AUG-1999;	99US-0147302P.
PR	05-AUG-1999;	99US-0147260P.
PR	06-AUG-1999;	99US-0147303P.
PR	09-AUG-1999;	99US-0147416P.
PR	09-AUG-1999;	99US-0147493P.
PR	10-AUG-1999;	99US-0147935P.
PR	11-AUG-1999;	99US-0148171P.
PR	12-AUG-1999;	99US-0148319P.
PR	13-AUG-1999;	99US-0148341P.
PR	13-AUG-1999;	99US-0148565P.
PR	16-AUG-1999;	99US-0148684P.
PR	17-AUG-1999;	99US-0149175P.
PR	18-AUG-1999;	99US-0149426P.
PR	20-AUG-1999;	99US-0149722P.
PR	20-AUG-1999;	99US-0149723P.
PR	20-AUG-1999;	99US-0149929P.
PR	23-AUG-1999;	99US-0149930P.
PR	25-AUG-1999;	99US-0150566P.
PR	26-AUG-1999;	99US-0150884P.
PR	27-AUG-1999;	99US-0151065P.
PR	27-AUG-1999;	99US-0151066P.
PR	30-AUG-1999;	99US-0151303P.
PR	31-AUG-1999;	99US-0151438P.
PR	01-SEP-1999;	99US-0151930P.
PR	07-SEP-1999;	99US-0152363P.
PR	10-SEP-1999;	99US-0153070P.
PR	13-SEP-1999;	99US-0153758P.
PR	15-SEP-1999;	99US-0154018P.
PR	16-SEP-1999;	99US-0154039P.
PR	20-SEP-1999;	99US-0154779P.
PR	22-SEP-1999;	99US-0155139P.
PR	23-SEP-1999;	99US-0155486P.
PR	24-SEP-1999;	99US-0155659P.
PR	28-SEP-1999;	99US-0156458P.
PR	29-SEP-1999;	99US-0156596P.

PR	04-OCT-1999;	99US-0157117P.
PR	05-OCT-1999;	99US-0157753P.
PR	06-OCT-1999;	99US-0157865P.
PR	07-OCT-1999;	99US-0158029P.
PR	08-OCT-1999;	99US-0158232P.
PR	12-OCT-1999;	99US-0158369P.
PR	13-OCT-1999;	99US-0159293P.
PR	13-OCT-1999;	99US-0159294P.
PR	13-OCT-1999;	99US-0159295P.
PR	14-OCT-1999;	99US-0159329P.
PR	14-OCT-1999;	99US-0159330P.
PR	14-OCT-1999;	99US-0159331P.
PR	14-OCT-1999;	99US-0159637P.
PR	14-OCT-1999;	99US-0159638P.
PR	18-OCT-1999;	99US-0159584P.
PR	21-OCT-1999;	99US-0160741P.
PR	21-OCT-1999;	99US-0160767P.
PR	21-OCT-1999;	99US-0160770P.
PR	21-OCT-1999;	99US-0160814P.
PR	21-OCT-1999;	99US-0160815P.
PR	22-OCT-1999;	99US-0160980P.
PR	22-OCT-1999;	99US-0160981P.
PR	25-OCT-1999;	99US-0161404P.
PR	25-OCT-1999;	99US-0161405P.
PR	25-OCT-1999;	99US-0161406P.
PR	26-OCT-1999;	99US-0161359P.
PR	26-OCT-1999;	99US-0161360P.
PR	26-OCT-1999;	99US-0161361P.
PR	28-OCT-1999;	99US-0161920P.
PR	28-OCT-1999;	99US-0161992P.
PR	28-OCT-1999;	99US-0161993P.
PR	29-OCT-1999;	99US-0162142P.

Alignment Scores:

Pred. No.:

1.41e-48

Score:

602.00

Percent Similarity:

71.08%

Best Local Similarity:

54.89%

Query Match:

39.45%

DB:

3

Length:

1095

Matches:

129

Conservative:

38

Mismatches:

58

Indels:

10

Gaps:

5

US-10-660-226-10 (1-305) x AAC47100 (1-1095)

QY	57	ArgLeuArgAla-----LysLysSerSerGlyGlyHisGluAsnSer-----	70
Db	209	CGACTTAGATCAGTTTCTGATAAGAACTCCTCAGCATTGTTGGAAACTGGAAGTCTTCTT	268
QY	71	HisAsnSerValAspGlu-----AlaLeuLeuLeuLysArgLysSerGluGluValLeu	88
Db	269	CATTCTCCATTTGATGAAGAACACAGATTTTGAAGCAGAAAAAGCTGAGAGGTAAA	328
QY	89	PheTyrLeuAsnGlyArgCysIleTyrLeuValGlyMetMetGlySerGlyLysSerThr	108
Db	329	CCGTATTTAAATGGACGATCGATGATCTTGTGTGATGATGGTTCGGGAAAAACGACT	388
QY	109	ValGlyLysIleMetSerGluValleuGlyTyrSerPhePheAspSerAspLysLeuVal	128
Db	389	GTAGGGAAGATTATGCGAAGATCGCTTGGTTATACATTCTTTGATTGTGACACATTGTATC	448
QY	129	GluGlnAlaValGlyMetProSerValAlaGlnIlePheLysValHisSerGluAlaPhe	148
Db	449	GAGCAGGCTTATGAAGGAACTTCTGTAGCTGAGATATTTGACANTTTCGGTGAGAGTGC	508
QY	149	PheArgAspAsnGluSerSerValLeuArgAspLeuSerSerMet---ArgArgLeuVal	167
Db	509	TTTCAGAAAAAGAGACTGAACGGTTAAAGAAACTCTCTTTGATGTACCACCAAGTTGTT	568
QY	168	ValAlaThrGlyGlyAlaValleArgProIleAsnTyrArgTyrMetLysArgGly	187
Db	569	GTTCACCGGGGGGGCGAGTTATAGACCCCACTTATGAAGTACATTCATATAAAGT	628

PR	02-AUG-1999;	99US-0146389P.
PR	03-AUG-1999;	99US-0147038P.
PR	04-AUG-1999;	99US-0147204P.
PR	04-AUG-1999;	99US-0147302P.
PR	05-AUG-1999;	99US-0147192P.
PR	05-AUG-1999;	99US-0147260P.
PR	06-AUG-1999;	99US-0147303P.
PR	06-AUG-1999;	99US-0147416P.
PR	09-AUG-1999;	99US-0147493P.
PR	09-AUG-1999;	99US-0147935P.
PR	10-AUG-1999;	99US-0148171P.
PR	11-AUG-1999;	99US-0148319P.
PR	12-AUG-1999;	99US-0148341P.
PR	13-AUG-1999;	99US-0148565P.
PR	13-AUG-1999;	99US-0148684P.
PR	16-AUG-1999;	99US-0149368P.
PR	17-AUG-1999;	99US-0149175P.
PR	18-AUG-1999;	99US-0149426P.
PR	20-AUG-1999;	99US-0149722P.
PR	20-AUG-1999;	99US-0149723P.
PR	20-AUG-1999;	99US-0149929P.
PR	23-AUG-1999;	99US-0149902P.
PR	23-AUG-1999;	99US-0149930P.
PR	25-AUG-1999;	99US-0150566P.
PR	26-AUG-1999;	99US-0150884P.
PR	27-AUG-1999;	99US-0151065P.
PR	27-AUG-1999;	99US-0151066P.
PR	27-AUG-1999;	99US-0151080P.
PR	30-AUG-1999;	99US-0151303P.
PR	31-AUG-1999;	99US-0151438P.
PR	01-SEP-1999;	99US-0151930P.
PR	07-SEP-1999;	99US-0152363P.
PR	10-SEP-1999;	99US-0153070P.
PR	13-SEP-1999;	99US-0153758P.
PR	15-SEP-1999;	99US-0154018P.
PR	16-SEP-1999;	99US-0154039P.
PR	20-SEP-1999;	99US-0154779P.
PR	22-SEP-1999;	99US-0155139P.
PR	23-SEP-1999;	99US-0155486P.
PR	24-SEP-1999;	99US-0155659P.
PR	28-SEP-1999;	99US-0156458P.
PR	29-SEP-1999;	99US-0156596P.
PR	04-OCT-1999;	99US-0157117P.
PR	05-OCT-1999;	99US-0157753P.
PR	06-OCT-1999;	99US-0157865P.
PR	07-OCT-1999;	99US-0158029P.
PR	08-OCT-1999;	99US-0158232P.
PR	12-OCT-1999;	99US-0158369P.
PR	13-OCT-1999;	99US-0159293P.
PR	13-OCT-1999;	99US-0159294P.
PR	13-OCT-1999;	99US-0159295P.
PR	14-OCT-1999;	99US-0159329P.
PR	14-OCT-1999;	99US-0159330P.
PR	14-OCT-1999;	99US-0159331P.
PR	14-OCT-1999;	99US-0159637P.
PR	14-OCT-1999;	99US-0159638P.
PR	18-OCT-1999;	99US-0159584P.
PR	21-OCT-1999;	99US-0160741P.
PR	21-OCT-1999;	99US-0160767P.
PR	21-OCT-1999;	99US-0160768P.
PR	21-OCT-1999;	99US-0160770P.
PR	21-OCT-1999;	99US-0160814P.
PR	21-OCT-1999;	99US-0160815P.
PR	22-OCT-1999;	99US-0160980P.
PR	22-OCT-1999;	99US-0160981P.
PR	22-OCT-1999;	99US-0160989P.
PR	25-OCT-1999;	99US-0161404P.
PR	25-OCT-1999;	99US-0161405P.
PR	25-OCT-1999;	99US-0161406P.
PR	26-OCT-1999;	99US-0161359P.
PR	26-OCT-1999;	99US-0161360P.
PR	26-OCT-1999;	99US-0161361P.
PR	28-OCT-1999;	99US-0161920P.
<hr/>		
PR	28-OCT-1999;	99US-0161992P.
PR	28-OCT-1999;	99US-0161993P.
PR	29-OCT-1999;	99US-0162142P.
<hr/>		
Alignment Scores:		
Pred. No.:	4,146-46	Length: 1093
Score:	576.50	Matches: 129
Percent Similarity:	70.51%	Conservative: 36
Best Local Similarity:	55.13%	Mismatches: 60
Query Match:	37.78%	Indels: 10
DB:	3	Gaps: 5
<hr/>		
US-10-660-226-10 (1-305) x AAC36746 (1-1093)		
QY	57 ArgLeuArgAla-----LysLysSerSerGlyHisGluasnSer-----	70
DB	209 CGACTTAGATCAGTTTCTGTGATAAGAACTCCTCAGCATTTGTTGAAACTGGAGTCTTCTT	268
QY	71 HisAsnSerValAsp---GluAlaLeuLeuLysArgLysSerGluGluValLeuPhe	89
DB	269 CATTCTCCATTTGATGAAGAACACAGATTTTGAGAGAAAAGCTGAAGAGGTTAAACCG	328
QY	90 TyrLeuAsnGlyArgCysIleTyrLeuValGlyMetMetGlySerGlyLysSerThrVal	109
DB	329 TATTTAAATGGACGATCGATGTATCTTGTGTATGATGGGTTCGGGAAACG-ACGTGA	387
QY	110 GlyIysIleMetSerGluValLeuGlyTyrSerPhePheAspSerAspLysLeuValGlu	129
DB	388 GGGAAAGATTATGGCAAGATCGCTTGGTTATACATTTCTTGATTGTGACACTTTGATCGAG	447
QY	130 GlnAlaValGlyMetProSerValAlaGlnIlePheLysValHisSerGluAlaPhePhe	149
DB	448 CAGGTATGAAGGGAACTTCTGTAGTCAGATATTTGACATTTTCGGTGAGAGTCTCTC	507
QY	150 ArgAspAsnGluSerSerValLeuArgAspLeuSerSerMet---ArgArgLeuValVal	168
DB	508 AGAGAAAAAGAGACTGAAGCGTTAAAGAAACTCTCTTTGATGTACCAACCAAGTTGTGT	567
QY	169 AlaThrGlyGlyAlaValIleArgProIleAsnTyrArgTyrMetLysArgGlyLeu	188
DB	568 TCACCGGGGAGGGCAGTTATAAGACCCATCAATTGGAAGTACATGATAAAGGTATT	627
QY	189 SerValTyrLeuAspValProLeuAspAlaLeuAlaArgArgIleAlaLysValGlyThr	208
DB	628 AGTATTGGCTTGATGATACCTCTAGAGCCTTAGCGCATAGATAGTCTCTAGGAACT	687
QY	209 AlaSerArgProLeuLeu---AspGlnProSerGlyAspProTyrAlaMetAlaPheSer	227
DB	688 GGTTCAGACCACTTGTACATGATGATGAGTCAGGCGGACACATACACAGCGCTTTAAAC	747
QY	228 LysLeuSerMetLeuAlaGlnGlnArgGlyAspAlaTyrAlaAsnAlaAspValArgVal	247
DB	748 CGTCTTTCAACGATTTGGGATGCAGTGTGAGCATACACTAAAGCCAGCCAGAGTT	807
QY	248 SerLeuGluGluIleAlaCysLysGlnGlyHisAspAspValSerLysLeuThrProThr	267
DB	808 TCCTCGGAGATATTACTTTGAAGCTCGGTATAGAGTGTCTCAGATCTTACACAGCT	867
QY	268 AspIleAlaIleGluSerLeuHisLysIleGluSerPheVal	281
DB	868 GAAATCGCCATTTGAGCGCTTTGAGCAAGTTTCAGAGCTATTTA	909
<hr/>		
RESULT 9		
AAC41622		
ID	AAC41622 standard; DNA; 696 BP.	
XX	AAC41622;	
AC	AAC41622;	
XX	17-OCT-2000 (first entry)	
DT	Arabidopsis thaliana DNA fragment SEQ ID NO: 32537.	
XX	Hybridisation assay; genetic mapping; gene expression control;	
DE		
XX		
KW		

KW protein identification; signal transduction pathway; metabolic pathway;
KW promoter; termination sequence; ss.
XX Arabidopsis thaliana.
OS EP1033405-A2.
PN
XX
XX 06-SEP-2000.
PD
XX 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0130891P.
PR 30-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.

CC with the property, and growing the transformed plant. The polypeptide is
CC useful for improving plant cold tolerance, manipulating growth rate in
CC plant cells by modification of the cell cycle pathway, improving plant
CC drought tolerance, providing increased resistance to plant disease,
CC producing galactanman (or lignin or plant growth regulators), improving
CC plant heat tolerance, improving plant tolerance to herbicides, increasing
CC the rate of homologous recombination in plants, improving plant tolerance
CC to extreme osmotic conditions or to pathogens or pests, improving yield
CC by modification of photosynthesis, modifying seed oil or protein yield
CC and/or content, improving yield by modification of carbohydrate, nitrogen
CC or phosphorus use and/or uptake, or improving yield by providing improved
CC plant growth and development under at least one stress condition. The
CC polynucleotide and polypeptide may also be used in recombinant DNA
CC constructs, in physical arrays of molecules, as plant breeding markers,
CC or in computer-based storage and analysis systems. The present sequence
CC is a Cotton plant cDNA of the invention. NOTE: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?DocID=2004018130. However only 6585
CC polynucleotide sequences were available, the remaining 5213
CC polynucleotides and all 58798 protein sequences were not present.

XX Sequence 1010 BP; 286 A; 180 C; 234 G; 310 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 3.97e-36 Length: 1010
Score: 473.00 Matches: 121
Percent Similarity: 59.53% Conservative: 32
Best Local Similarity: 47.08% Mismatches: 61
Query Match: 31.00% Indels: 45
DB: 13 Gaps: 5

US-10-660-226-10 (1-305) x ADR60835 (1-1010)

QY 30 ThrGlySerLeuArgValAlaAspProAlaGlyProAlaValAlaValArgAlaGly 49
DB 214 ACTGGTCTCTTGGGTTTC-----GATCGGATAGCAAGCAACAAG 258
QY 50 SerLysProValAlaProLeuArgLeuArgAlaLysLysSerGlyGlyHisGlu--- 68
DB 259 ACCCGGGTGTGTTTTCGGCCCACTTTCGGTTTCAGACATCTTCAATGACATGATCA 318
QY 68 ----- 68
DB 319 GTCTCTTTGGAGGTTTCTGCTCTCTATACGAACCTTTTCAGCTTCAACAATGGAAC 378
QY 69 ---AsnSerHisAsnSerValAspGluAlaLeuLeuLeuLysArgLysSerGluGluVal 87
DB 379 GGCAGCATTCATGCACCTTTTGGTGAAGCTTTAATGTTAAAGAAATAAGTCACTAGAGTT 438
QY 88 LeuPheTyLeuAsnGlyArgCysIleTyLeuValGlyMetMetGlySerGlyLysSer 107
DB 439 GAGCCATATTAATGGGCACAGTATATATCTTTGGATGATGGGCTCTGGAACCT- 497
QY 108 ThrValGlyLysIleMetSerGluValLeuGlyTyzSerPhePheAspSerAspLysLeu 127
DB 498 ACAGTGGCAAAATCTCTCCCATGACTCGGTTATTCATTT-----TGTGACAGGTTG 551
QY 128 -ValGluGluAlaValGlyMetPro----- 135
DB 552 AATATTGAGAGTGTATTCTCTCTGATGTCTTACCAATTTCCAAGAGGTGAATGGAAT 611
QY 136 -SerValAlaGluIlePheLysValHisSerGluAlaPhePheArgAspAsnGluSerSe 155
DB 612 GTCTGTAGCTGGAATATTTAAGCTTATGCGGAGAGTTTCTCAGAAAAAAGAGACTGA 671
QY 155 rValLeuArgAspLysSerSerMetArgLeuValValAlaThrGlyGlyAlaVal 175
DB 672 GGTATTGCAGAGCTTCTTCACAGAAACAACATTTGTTTCTACTGGCGAGGTGCAGT 731
QY 175 lIleArgProIleAsnTrpArgTyMet----LysArgGlyLeuSerValTrpLeuAspVa 194
DB 732 TGTACAGGATGTGAACACTGGGACTATATGCAGAGAAGAGGATTTGTTCTCTGTTAGATGT 791

QY 194 lProLeuAspAlaLeuAlaArgIleAlaLysValGlyThrAlaSerArgProLeuLe 214
DB 792 ACCTTTGGAGCCTTGGCACAAAGGATGTCAGTAGTACTCATTTCTCGTCCCTTTT 851
QY 214 uAspGlnProSerGlyAspProTyTyAlaMetAlaPheSerLysLeuSerMetLeuAlaGl 234
DB 852 GCATTATGAAGATGCGGATCCATATACAAAGGCTTA-AAACGCTCTGTCTTACCTTTTGA 910
QY 234 nGlnArgGlyAspAlaTyAlaAsnAlaAspValArgValSerLeuGlu 250
DB 911 ACAGAGGGGTAAAAAATATGCTAAAGCAATGCCAGGGTTTCATTGGAA 959

RESULT 11

ADK58257

ID ADK58257 standard; DNA; 660 BP.

XX AC ADK58257;

XX DT 06-MAY-2004 (first entry)

XX DE Plant DNA sequence which confers altered metabolic characteristic #5640.

XX KW altered metabolic characteristic; plant; acid metabolism;

XX KW alcohol metabolism; fatty acid metabolism;

XX KW branched fatty acid metabolism; alkaloid metabolism;

XX KW amino acid metabolism; ester metabolism; glyceride metabolism;

XX KW phenolic metabolism; carbohydrate metabolism; sterol metabolism;

XX KW terpene metabolism; isoprenoid metabolism; alkene metabolism;

XX KW alkyne metabolism; hydrocarbon metabolism; ketone metabolism;

XX KW quinone metabolism; disease resistance; gene shuffling; sexual PCR; ds.

XX OS Unidentified.

XX PN WO2003020936-A1.

XX PD 13-MAR-2003.

XX PF 30-AUG-2002; 2002WO-US027884.

XX PR 31-AUG-2001; 2001US-0316471P.

XX PA (DOWC) DOW CHEM CO.

XX PA (DOWC) DOW AGROSCIENCES LLC.

XX PI Weglarz T, Gachotte D, Blakelee B, McCreary DA, Pell RJ;

XX PI Oriedo JVB, Crosley R, Reddy AS, Shukla V, Larrinua I, Miller BA;

XX DR WPI; 2003-313091/30.

XX PT Novel genes that confer altered metabolic characteristics in Nicotiana
PT benthamiana plants, useful for altering the levels of metabolites e.g.
PT acids, fatty acids, amino acids, carbohydrates, hydrocarbons and sterols.

XX PS Claim 1; SEQ ID NO 5640; 2576pp; English.

XX CC The invention comprises DNA sequences which confer an altered metabolic
CC characteristic when they are expressed in a plant. The DNA sequences of
CC the invention are useful for producing plants with an altered metabolic
CC characteristic, such as: altered acid metabolism, alcohol metabolism,
CC fatty acid metabolism, branched fatty acid metabolism, alkaloid or other
CC base metabolism, altered amino acid metabolism, altered ester metabolism,
CC altered glyceride metabolism, altered phenolic metabolism, altered
CC carbohydrate metabolism, altered sterol, oxygenated terpene, or
CC isoprenoid metabolism, alkene or alkyne metabolism, hydrocarbon
CC metabolism, ketone or quinone metabolism. The DNA sequences of the
CC invention may be used to provide disease resistance in a plant and gene
CC shuffling or sexual PCR procedures. The present nucleic acid represents a
CC DNA sequence of the invention.

XX SQ Sequence 660 BP; 194 A; 74 C; 168 G; 224 T; 0 U; 0 Other;

Alignment Scores:


```
Db 61 GCAAGAAGAAATGCTGCTAGGAACCGCGTCTCGACCACTCTTGATCAGGAATCCGGT 120
Qy 220 AspProTyAlaMetAlaPheSerLysLeuSerMetLeuAlaGlnGlnArgGlyAspAla 239
Db 121 GATCCTATGCAAGCTTATGCAAACTTAGCTACCTTTTGGAGCAAGATGACTCG 180
Qy 240 TyrAlaAsnAlaAspValArgValSerLeuGluGluLeuAlaCysGlnGlyHisAsp 259
Db 181 TATGCTAATGCTGATGCCAGAGTTTCACTTGAACATATTGCATTAAACAAGGCCATAAT 240
Qy 260 AspValSerLysLeuThrProThrAspPheAlaLeuGluSerLeuHisLysIleGluSer 279
Db 241 GATGTCATATACCTTACACCTAGTACCATCGCATTTGAGGCATTCGTAAGATGGGAAGT 300
Qy 280 PheVal 281
Db 301 TTTCCT 306

RESULT 13
ADK58255
ID ADK58255 standard; DNA; 329 BP.
AC ADK58255;
XX
XX
DT 06-MAY-2004 (first entry)
XX
DE Plant DNA sequence which confers altered metabolic characteristic #5638.
XX
KW altered metabolic characteristic; plant; acid metabolism;
KW alcohol metabolism; fatty acid metabolism;
KW branched fatty acid metabolism; alkaloid metabolism;
KW amino acid metabolism; ester metabolism; glyceride metabolism;
KW phenolic metabolism; carbohydrate metabolism; sterol metabolism;
KW terpene metabolism; isoprenoid metabolism; alkene metabolism;
KW alkyne metabolism; hydrocarbon metabolism; ketone metabolism;
KW quinone metabolism; disease resistance; gene shuffling; sexual PCR; ds.
XX
OS Unidentified.
XX
XX WO2003020936-A1.
XX
XX 13-MAR-2003.
XX
XX 30-AUG-2002; 2002WO-US027884.
XX
XX 31-AUG-2001; 2001US-0316471P.
XX
XX (DOWC ) DOW CHEM CO.
XX (DOWC ) DOW AGROSCIENCES LLC.
XX
XX Weglarz T, Gachotte D, Blakeslee B, McCreary DA, Pell RJ;
XX Oriedo JVB, Crosley R, Reddy AS, Shukla V, Larrinua I, Miller BA;
XX WPI; 2003-313091/30.
XX
XX Novel genes that confer altered metabolic characteristics in Nicotiana
XX benthamiana plants, useful for altering the levels of metabolites e.g.
XX acids, fatty acids, amino acids, carbohydrates, hydrocarbons and sterols.
XX
XX Claim 1; SEQ ID NO 5638; 2576pp; English.
XX
XX The invention comprises DNA sequences which confer an altered metabolic
XX characteristic when they are expressed in a plant. The DNA sequences of
XX the invention are useful for producing plants with an altered metabolic
XX characteristic, such as: altered acid metabolism, alcohol metabolism,
XX fatty acid metabolism, branched fatty acid metabolism, alkaloid or other
XX base metabolism, altered amino acid metabolism, altered ester metabolism,
XX altered glyceride metabolism, altered phenolic metabolism, altered
XX carbohydrate metabolism, altered sterol, oxygenated terpene, or
XX isoprenoid metabolism, alkene or alkyne metabolism, hydrocarbon
XX metabolism, ketone or quinone metabolism. The DNA sequences of the
XX invention may be used to provide disease resistance in a plant and gene
XX shuffling or sexual PCR procedures. The present nucleic acid represents a
```

```
CC DNA sequence of the invention.
XX
XX Sequence 329 BP; 93 A; 52 C; 88 G; 96 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 1,31e-26 Length: 329
Score: 368.00 Matches: 69
Percent Similarity: 82.69% Conservative: 17
Best Local Similarity: 66.35% Mismatches: 18
Query Match: 24.12% Indels: 0
DB: 10 Gaps: 0
US-10-660-226-10 (1-305) x ADK58255 (1-329)
Qy 91 LeuAsnGlyArgCysIleTyrLeuValGlyMetMetGlySerGlyLysSerThrValGly 110
Db 16 TTGAATGGACGATCTATGTACTCTTGTGGAATGATGGGTTCTGGGAAAAACAACCTGGGA 75
Qy 111 LysIleMetSerGluValLeuGlyTyrSerPhePheAspSerAspLysLeuValGluGln 130
Db 76 AAGTTAATGTCCAAAGTCTCGGTTATACGTTCTTTGACTCGGACACCTTTGATTGAACAG 135
Qy 131 AlaValGlyMetProSerValAlaGlnIlePheLysValHisSerGluAlaPhePheArg 150
Db 136 GCGATGAATGGAACCTTCGTTGCGAGATATTTGTTTCATCAGCGAGAAATTTTGTAGA 195
Qy 151 AspAsnGluSerSerValLeuArgAspLeuSerMetArgArgLeuValValAlaThr 170
Db 196 GGAAAGGAGACCGATCGCTTTAAGAAGCTCTCTTCGAGGTATCAAGTTGTTGTTCCACA 255
Qy 171 GlyGlyGlyAlaValIleArgProIleAsnTrpArgTyrMetLysArgGlyLeuSerVal 190
Db 256 GGTGAGGTGCAGTTTATAAGACCCCATTAACCTGGAAGTATATGCAATAAAGGAATCAGCATT 315
Qy 191 TrpLeuAspVal 194
Db 316 TGGCTAGATGTG 327

RESULT 14
ADR60836
ID ADR60836 standard; cDNA; 614 BP.
XX
XX ADR60836;
XX
DT 02-DEC-2004 (first entry)
XX
XX Cotton cDNA sequence, SEQ ID 1617.
XX
KW Cotton; ss; plant; cold tolerance; growth rate; cell cycle pathway;
KW drought tolerance; plant disease resistance; galactomannan; lignin;
KW plant growth regulator; heat tolerance; herbicide tolerance;
KW homologous recombination; extreme osmotic condition tolerance;
KW pathogen resistance; pest resistance; yield; photosynthesis; seed oil;
KW stress resistance.
XX
XX Gossypium hirsutum.
XX
XX US2004181830-A1.
XX
XX 16-SEP-2004.
XX
XX 29-JAN-2004; 2004US-00767795.
XX
XX 07-MAY-2001; 2001US-00849529.
XX 12-DEC-2001; 2001US-00021323.
XX
XX (KOVA/) KOVALIC D K.
XX (ZHOU/) ZHOU Y.
XX (CAOY/) CAO Y.
XX
XX Kovalic DK, Zhou Y, Cao Y;
XX WPI; 2004-667718/65.
XX
XX
```

XX New recombinant nucleic acid molecules and polypeptides from *Gossypium*
PT hirsutum, useful for producing plants with improved biological
PT characteristics (e.g. improved plant cold or drought tolerance).
XX
XX
XX Claim 1; SEQ ID NO 1617; 14pp; English.
XX
XX The invention relates to a recombinant polynucleotide comprising any of
CC the 58798 Cotton plant cDNA sequences mentioned in the specification.
CC Also a recombinant polypeptide comprising any of the 58798 amino acid
CC sequences mentioned in the specification and producing a plant having an
CC improved property. Producing a plant having an improved property
CC comprises transforming a plant with a recombinant construct comprising a
CC promoter region functional in a plant cell operably joined to a
CC polynucleotide comprising a coding sequence for a polypeptide associated
CC with the property, and growing the transformed plant. The polypeptide is
CC useful for improving plant cold tolerance, manipulating growth rate in
CC plant cells by modification of the cell cycle pathway, improving plant
CC drought tolerance, providing increased resistance to plant disease,
CC producing galactanmannan (or lignin or plant growth regulators), improving
CC plant heat tolerance, improving plant tolerance to herbicides, increasing
CC the rate of homologous recombination in plants, improving plant tolerance
CC to extreme osmotic conditions or to pathogens or pests, improving yield
CC by modification of photosynthesis, modifying seed oil or protein yield
CC and/or content, improving yield by modification of carbohydrate, nitrogen
CC or phosphorus use and/or uptake, or improving yield by providing improved
CC plant growth and development under at least one stress condition. The
CC polynucleotide and polypeptide may also be used in recombinant DNA
CC constructs, in physical arrays of molecules, as plant breeding markers,
CC or in computer-based storage and analysis systems. The present sequence
CC is a Cotton plant cDNA of the invention. NOTE: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?DocID=20040181830. However only 6585
CC polynucleotide sequences were available, the remaining 52213
CC polynucleotides and all 58798 protein sequences were not present.
XX
SQ Sequence 614 BP; 170 A; 98 C; 154 G; 192 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1,04e-25 Length: 614
Score: 362.50 Matches: 80
Percent Similarity: 73.72% Conservative: 21
Best Local Similarity: 58.39% Mismatches: 34
Query Match: 23.75% Indels: 2
DB: 13 Gaps: 1

US-10-660-226-10 (1-305) x ADR60836 (1-614)

QY 149 PheArgAspAsnGluSerSerValLeuArgAspLeuSerSerMetArgLeuValVal 168
DB 1 TTCAGAAAGAGAGACTGAGGTATTGCAGAGGCTCTTCAAGAAACAGCTGTGTGT 60
QY 169 AlaThrGlyGlyAlaValIleArgProIleAsnTrpArgTyrMet---LysArgGly 187
DB 61 TCTACTGGCGAGGTGCAGTGTACGGGTGTGACTGGACTATATGCAGAAAGAGGG 120
QY 188 LeuSerValTrpLeuAspValProLeuAspAlaLeuAlaArgArgIleAlaLysValGly 207
DB 121 GTTGTGTCTGTTAGATTGATCCCTTTGGAAGCCTTGGCACAGAGGATTCCTCAGTAGGT 180
QY 208 ThrAlaSerArgProLeuLeuAspGlnProSerGlyAspProTyrAlaMetAlaPheSer 227
DB 181 ACTCATCTCGTCCCTTTTGCAATATGAACTATGACATGGCGATCCCTATCAAGAGCTTTAAA 240
QY 228 LysLeuSerMetLeuAlaGlnGlnArgGlyAspAlaTyrAlaAsnAlaAspValArgVal 247
DB 241 CGGCTGTCTTACCTTTTGGAGCTGAGGGGTAAATAATTATGCTAAAGCAATGCCGGTT 300
QY 248 SerLeu-GluGluIleAlaCysLysGlnGlyHisAspValSerLysLeuThrProThr 267
DB 301 TCATTGTAAGAAGAAATGCTGGCTTACTAGGTATTATGAGATGTATCAGATCTTACTCCAC 360

QY 267 rAspIleAlaIleGluSerLeuHisValIleGluSerPheValIleGlu 283
DB 361 AGAGATCCCAATCGAGGCTTGGGCACAAATGGAGTGTATCTAAAGGAG 409
RESULT 15
ID ACN52580/c
XX ACN52580 standard; cDNA; 573 BP.
XX ACN52580;
XX 02-DEC-2004 (first entry)
XX Cotton androecium tissue EST Clone ID: LIB3828-016-Q1-N6-D8, SEQ:7361.
XX Cotton; plant; EST; expressed sequence tag; transgenic plant; androecium;
KW variety Nucotton33B; library LIB3828; molecular tag; molecular marker;
KW genetic mapping; molecular mapping; seed germination; plant growth;
KW plant quality; plant yield; plant breeding; tissue printing; ss.
XX *Gossypium hirsutum*.
XX US2004123340-A1.
XX 24-JUN-2004.
XX 12-DEC-2001; 2001US-00021323.
XX 14-DEC-2000; 2000US-0255619P.
XX (DEIK/) DEIKMAN J.
XX (FENG/) FENG P C C.
XX (FINC/) FINCHER K L.
XX (ZIEG/) ZIEGLER T E.
XX Deikman J, Feng PCC, Fincher KL, Ziegler TE;
XX WPI; 2004-479808/45.
XX
XX New isolated nucleic acid molecule that encodes a plant protein or its
PT fragment, useful for isolating a variety of agronomically significant
PT genes associated with plant growth, quality or yield, and as molecular
XX tags to map genes.
XX Claim 1; SEQ ID NO 7361; 34pp; English.
XX
XX The invention relates to 17880 cotton expressed sequence tags (ESTs;
CC ACN45220-ACN63099). The ESTs were isolated from cDNA libraries generated
CC from primed or non-primed seeds from variety DP50B, mature seeds from
CC variety Coker 312 Boswell 96 Field, and androecium tissue, gynoeceum
CC tissue, developing fibres, carpel walls and septa from variety
CC Nucotton33B. The invention also relates to substantially purified
CC proteins or their fragments encoded by nucleic acid molecules of the
CC invention, and to transformed plants having a nucleic acid construct
CC comprising a nucleic acid of the invention. The cotton ESTs are useful as
CC molecular tags to isolate genetic regions, to isolate genes, to map
CC genes, to determine gene function and to determine whether genes are
CC members of a particular gene family. The nucleic acid molecules may be
CC used for isolating a variety of agronomically significant genes
CC associated with plant growth, quality, yield, and could also serve as
CC links in metabolic and catabolic pathways. The nucleic acid molecules are
CC also useful for identifying genes important in initiating and maintaining
CC seed germination or that may be used to mitigate stresses encountered
CC during seed germination. The ESTs additionally enable the acquisition of
CC promoters and cis-regulatory elements which will be useful to express
CC agronomically significant genes in these tissues and/or other tissues,
CC and also permits the acquisition of molecular markers useful in breeding
CC schemes, genetic and molecular mapping, and in cloning of agronomically
CC significant genes. The nucleic acid molecules are further useful for
CC detecting the expression level or pattern of a protein or mRNA and for
CC detecting the presence or quantity of a protein by tissue printing. The
CC present sequence represents a specifically claimed EST isolated from a
CC cotton variety Nucotton33B androecium tissue cDNA library (LIB3828). The
CC sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from the US
XX patent office at seqdata.uspto.gov/sequence.html?DocID=US20040123340
SQ Sequence 573 BP; 183 A; 138 C; 86 G; 166 T; 0 U; 0 Other;

Alignment Scores:		2.68e-24	Length:	573
Pred. No.:		347.50	Matches:	73
Score:		75.21%	Conservative:	18
Percent Similarity:		60.33%	Mismatches:	29
Best Local Similarity:		22.77%	Indels:	1
Query Match:		13	Gaps:	1
DB:				
US-10-660-226-10 (1-305) x ACN52580 (1-573)				
QY	164	ArgArgLeuValAlaThrGlyGlyGlyValAlaValIleArgProIleAsnTrpArgTyr	183	
DB	568	AAACAACCTTGTCTTCTACTGTGGAGGTGCAGTTGTACTGGATGTGAACCTGGGACTAT	509	
QY	184	Met---LysArgGlyLeuSerValTrpLeuAspValProLeuAspAlaLeuAlaArgArg	202	
DB	508	ATGCAGAAGAAGGGGATTTGTTGGTTAGATGTACCTTTGGAAGCCTTGGCACAAGG	449	
QY	203	IleAlaIysValGlyThrAlaSerArgProLeuLeuAspGlnProSerGlyAspProTyr	222	
DB	448	ATTGCTGCAGTAGGTACTCATCTCGTCCCTTTTGCATTATGAGATGGTGATCCATAT	389	
QY	223	AlaMetAlaIlePheSerLysLeuSerMetLeuAlaGlnGlnArgGlyAspAlaTyrAlaAsn	242	
DB	388	ACAAAGCCTTAAACGTCGTCTTACCTTTTGGAGCAGAGGGTAAAAATTATGCTAAA	329	
QY	243	AlaAspValArgValSerLeuGluGluIleAlaCysLysGlnGlyHisAspAspValSer	262	
DB	328	GCAAATGCCAGGCTTTTCATTGTAAAGAAATTGCTGGCAAACTAGGTTATAGAGATGTATCA	269	
QY	263	LysLeuThrProThrAspIleAlaIleGluSerLeuHisLysIleGluSerPheValIle	282	
DB	268	GATCTTACTCCAAACAGAGATCGCAATCGAGGCATTGGGACAAATTGGAGTGTATCTAAAG	209	
QY	283	Glu 283		
DB	208	GAG 206		

Search completed: August 25, 2005, 20:53:50
Job time : 618 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 25, 2005, 20:28:39 ; Search time 203 Seconds

(without alignments)
2458.445 Million cell updates/sec

Title: US-10-660-226-10

Perfect score: 1526

Sequence: 1 MEAGGVGLALQARAAGFGSS.....ADSSASDAQESQIQRIQTL 305

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DRV=xlh
-Q=/cgn2_1/USPTO spoal/US1060226/runat_25082005_110428_28555/app_query.fasta_1.455
-DB=Issued Patents NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human4.0.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US1060226 @CNC 1.1.69 @runat_25082005_110428_28555 -NCPU=6 -ICPU=3
-NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DRV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents NA: *
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq: *
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq: *
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq: *
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq: *
5: /cgn2_6/ptodata/1/ina/PTCUS_COMB.seq: *
6: /cgn2_6/ptodata/1/ina/backfiles1.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	245	16.1	696	4	US-09-540-236-167
2	245	16.1	14335	4	US-09-596-002-11
3	222	14.5	603	4	US-09-328-352-2023
4	222	14.5	603	4	US-09-489-039A-3641
5	214	14.0	1830121	4	US-09-557-884-1
6	214	14.0	1830121	4	US-09-643-990A-1
7	212.5	13.9	576	4	US-09-543-681A-2510
8	210	13.8	640681	4	US-09-790-988-1
9	205.5	13.5	555	4	US-09-252-991A-16393
10	205.5	13.5	576	4	US-09-252-991A-16202
11	185.5	12.2	537	4	US-09-107-532A-2146
12	179	11.7	477	4	US-09-583-110-886

c 13	179	11.7	480	4	US-09-107-433-795	Sequence 795, App
c 14	177	11.6	19390	3	US-08-961-527-86	Sequence 86, Appl
c 15	167.5	11.0	513	4	US-09-134-000C-1574	Sequence 1574, Ap
c 16	164.5	10.8	4403765	3	US-09-103-840A-2	Sequence 2, Appli
c 17	164.5	10.8	4411529	3	US-09-103-840A-1	Sequence 1, Appli
c 18	162	10.6	534	4	US-09-543-681A-1141	Sequence 1141, Ap
c 19	150	9.8	643	3	US-09-064-693A-21	Sequence 21, Appl
c 20	150	9.8	4530	3	US-09-064-693A-26	Sequence 26, Appl
c 21	145	9.5	642	4	US-09-489-039A-6367	Sequence 6367, Ap
c 22	143.5	9.4	1230025	4	US-09-198-452A-1	Sequence 1, Appli
c 23	143.5	9.4	1230230	4	US-09-438-185A-1	Sequence 1, Appli
c 24	140.5	9.2	519	3	US-09-134-001C-773	Sequence 773, App
c 25	139	9.1	2306	4	US-09-620-312D-166	Sequence 166, App
c 26	125	8.2	828	4	US-09-248-796A-4092	Sequence 4092, Ap
c 27	125	8.2	4656	3	US-09-425-665-1	Sequence 1, Appli
c 28	125	8.2	4656	3	US-09-685-668-1	Sequence 1, Appli
c 29	123.5	8.1	8614	4	US-09-902-540-787	Sequence 787, App
c 30	118.5	7.8	3150	4	US-09-710-279-3698	Sequence 3698, Ap
c 31	109	7.1	19580	2	US-08-500-857A-3	Sequence 3, Appli
c 32	108	7.1	18598	4	US-09-902-540-1143	Sequence 1143, Ap
c 33	106.5	7.0	2234	4	US-09-902-540-280	Sequence 280, App
c 34	105.5	6.9	8320	4	US-09-902-540-913	Sequence 913, App
c 35	105	6.9	1158	4	US-09-252-991A-1379	Sequence 1379, Ap
c 36	104.5	6.8	14570	4	US-09-902-540-1012	Sequence 1012, Ap
c 37	104	6.8	3150	4	US-09-252-991A-8765	Sequence 8765, Ap
c 38	103.5	6.8	5312	4	US-09-902-540-4775	Sequence 4775, Ap
c 39	103.5	6.8	27903	4	US-09-902-540-1235	Sequence 1235, Ap
c 40	103	6.7	2571	4	US-09-252-991A-9022	Sequence 9022, Ap
c 41	102.5	6.7	6819	4	US-09-032-438C-2	Sequence 2, Appli
c 42	102.5	6.7	9091	4	US-09-902-540-949	Sequence 949, App
c 43	102.5	6.7	10301	4	US-09-902-540-985	Sequence 985, App
c 44	102.5	6.7	35408	3	US-08-973-334-3	Sequence 3, Appli
c 45	102.5	6.7	35408	3	US-09-563-869A-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1

US-09-540-236-167
; Sequence 167, Application US/09540236
; Patent No. 6673910
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATAP
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2005-001
; CURRENT APPLICATION NUMBER: US/09/540,236
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 3840
; SEQ ID NO 167
; LENGTH: 696
; TYPE: DNA
; ORGANISM: M.catarrhalis
US-09-540-236-167

Alignment Scores:
Pred. No.: 4.71e-17 Length: 696
Score: 245.00 Matches: 73
Percent Similarity: 52.42% Conservative: 46
Best Local Similarity: 32.16% Mismatches: 76
Query Match: 16.06% Indels: 32
DB: 4 Gaps: 8

US-10-660-226-10 (1-305) x US-09-540-236-167 (1-696)

QY	78	LeuLeuLeuLysArg-----LysSerGluGluValLeuPheTyLeuAnGly-ArgCy	95
Db	80	CTTTCATTGAAGAGGTGATGTCGAAAAAGCAGGTAATGCCTTATCAAAACAACTACCTG	139
QY	95	e-IleTyLeuValGlyMetMetGlySerGlyLysSerThrValGlyLysIleMetSerG	115
Db	140	CGATTTTGTGGGACCTATGTTGGTCAGGCAAAACGATAGGAAGCTGCTTGCCA	199


```
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US 60/074,788
; CURRENT FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 16393
; LENGTH: 555
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-16393

Alignment Scores:
Pred. No.: 6,04e-13 Length: 555
Score: 205.50 Matches: 69
Percent Similarity: 49.29% Conservatve: 35
Best Local Similarity: 32.70% Mismatches: 71
Query Match: 13.47% Indels: 37
DB: 4 Gaps: 8

US-10-660-226-10 (1-305) x US-09-252-991A-16393 (1-555)
QY 66 GlyHisGluAAsn-----SerHisAAsnSerValAspGluAlaLeuLeuLys 81
Db 4 GGCCATCGCAATCGGTCGCTGATATCGAACACAGTGGTCAAC-----45
QY 82 ArgLysSerGluValLeuPheTyrLeuAAsnGlyArgCysIleTyrLeuValGlyMet 101
Db 46 -----CTGATCTCTCGCGCCCG 63
QY 102 MetGlySerGlyLysSerThrValGlyLysIleMetSerGluValLeuGlyTyrSerPhe 121
Db 64 ATGGGTGCTGGAAGAGACACCATCGGCGCTGCTCGCAAGGAGTGTCTCGCGTTC 123
QY 122 PheAspSerAspLysLeuValGluGlnAlaValGlyMetProSerValAlaGlnIlePhe 141
Db 124 AAGGACTCCGACAAAGAGATCGAACACAGCGGTGCGGC---GCGAATATCCCATGATCTTC 180
QY 142 LysValHisSerGluAlaPheArgAspAAsnGluSerSerValLeuArgAspLeuSer 161
Db 181 GATGTCAGGGCGAGTCCGTTCCCGAGCGGACGAGCCATGCTCACCGAATCTCG 240
QY 162 SerMetArgLysValValAlaThrGlyGlyAlaValIleArgProIleAsnTrp 181
Db 241 GCGCGCGACGGCATGTGATCGCCACCGCGCGCGGGCGGTGATGCGCAGCGTAACCG 300
QY 182 ArgTyrMetLysArg---GlyLeuSerValTrpLeuAspValProLeuAspAlaLeuAla 200
Db 301 CAGGTCTCGCGCGCGGTGCGCGGGTGTGTACTGCTGATGCTCGGTGAGCACCAGATC 360
QY 201 ArgArgIleAlaLysValGlyThrAlaSerArgProLeuLeuAspGlnProSerGlyAsp 220
Db 361 GCGCGCAGCGCGCGG-----GACCGCAACCGCCCCCTTGTGAGAAAGCCCAACCGCGGA 414
QY 221 ProTyrAlaMetAlaPheSerLysLeuSerMetLeuAlaGlnGlnArgGlyAspAlaTyr 240
Db 415 CAGATC-----CTCCGCGACTGATGGCGGTGCGGATCGGCTCTAC 456
QY 241 AlaAsn---AlaAspValArgValSerLeuGluGluIleAlaCysLysGlnGlyHisAsp 259
Db 457 CGCGAAATCGCGCATGTGTGTGGAGACCGAGCAAGG-GCCGCCCGCGCTGTCGTCTCCA 515
QY 260 AspValSerLysLeuThrProThrAspIleAla 270
Db 516 GGAATTTCTCGA---ACGCCCTGCGCAAGTTGCC 545

RESULT 10
US-09-252-991A-16202/c
; Sequence 16202, Application US/0925291A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
```

```
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 16202
; LENGTH: 576
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-16202

Alignment Scores:
Pred. No.: 6,38e-13 Length: 576
Score: 205.50 Matches: 69
Percent Similarity: 49.29% Conservatve: 35
Best Local Similarity: 32.70% Mismatches: 71
Query Match: 13.47% Indels: 37
DB: 4 Gaps: 8

US-10-660-226-10 (1-305) x US-09-252-991A-16202 (1-576)
QY 66 GlyHisGluAAsn-----SerHisAAsnSerValAspGluAlaLeuLeuLys 81
Db 573 GGCCATCGCAATCGGTCGCTGATATCGAACACAGTGGTCAAC-----532
QY 82 ArgLysSerGluValLeuPheTyrLeuAAsnGlyArgCysIleTyrLeuValGlyMet 101
Db 531 -----CTGATCTCTCGCGCCCG 514
QY 102 MetGlySerGlyLysSerThrValGlyLysIleMetSerGluValLeuGlyTyrSerPhe 121
Db 513 ATGGGTGCTGGAAGAGACACCATCGGCGCTGCTCGCAAGGAGTGTCTCGCGTTC 454
QY 122 PheAspSerAspLysLeuValGluGlnAlaValGlyMetProSerValAlaGlnIlePhe 141
Db 453 AAGGACTCCGACAAAGAGATCGAACACAGCGGTGCGGC---GCGAATATCCCATGATCTTC 397
QY 142 LysValHisSerGluAlaPheArgAspAAsnGluSerSerValLeuArgAspLeuSer 161
Db 396 GATGTCAGGGCGAGTCCGTTTCGCGAGCGGACGAGCCATGCTCACCGAATCTTC 337
QY 162 SerMetArgLysValValAlaThrGlyGlyAlaValIleArgProIleAsnTrp 181
Db 336 GCGCGCGACGGCATGTGATCGCCACCGCGCGGGCGGTGATGCGCAGCGTAACCG 277
QY 182 ArgTyrMetLysArg---GlyLeuSerValTrpLeuAspValProLeuAspAlaLeuAla 200
Db 276 CAGGTCTCGCGCGCGGTGCGCGGGTGTGTACTGCTGATGCTCGGTGAGCACCAGATC 217
QY 201 ArgArgIleAlaLysValGlyThrAlaSerArgProLeuLeuAspGlnProSerGlyAsp 220
Db 216 GCGCGCAGCGCGCGG-----GACCGCAACCGCCCCCTTGTGAGAAAGCCCAACCGCGGA 163
QY 221 ProTyrAlaMetAlaPheSerLysLeuSerMetLeuAlaGlnGlnArgGlyAspAlaTyr 240
Db 162 CAGATC-----CTCCGCGACTGATGGCGGTGCGGATCGGCTCTAC 121
QY 241 AlaAsn---AlaAspValArgValSerLeuGluGluIleAlaCysLysGlnGlyHisAsp 259
Db 120 CGCGAAATCGCGCATGTGTGTGGAGACCGAGCAAGG-GCCGCCCGCGCTGTCGTCTCCA 62
QY 260 AspValSerLysLeuThrProThrAspIleAla 270
Db 61 GGAATTTCTCGA---ACGCCCTGCGCAAGTTGCC 32

RESULT 11
US-09-107-532A-2146
; Sequence 2146, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
```

APPLICANT: Lynn A Doucette-Stamm and David Bush
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
NUMBER OF SEQUENCES: 7310
CORRESPONDENCE ADDRESS:
ADDRESSEE: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354

COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: PC
OPERATING SYSTEM: <Unknown>
SOFTWARE: ASCII

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,532A
FILING DATE: 30-Jun-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/085,598
FILING DATE: 14 May 1998
APPLICATION NUMBER: 60/051571
FILING DATE: July 2, 1997

ATTORNEY/AGENT INFORMATION:
NAME: Ariniello, Pamela Deneke
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-012
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 2146:
SEQUENCE CHARACTERISTICS:
LENGTH: 537 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO

ORIGINAL SOURCE:
ORGANISM: Enterococcus faecium
FEATURE:
NAME/KEY: misc feature
LOCATION: (B) LOCATION 1...537
SEQUENCE DESCRIPTION: SEQ ID NO: 2146:
US-09-107-532A-2146

Alignment Scores:
Pred. No.: 8,19e-11 Length: 537
Score: 185.50 Matches: 44
Percent Similarity: 61.34% Conservative: 29
Best Local Similarity: 36.97% Mismatches: 43
Query Match: 12.16% Indels: 3
DB: 4 Gaps: 3

US-10-660-226-10 (1-305) x US-09-107-532A-2146 (1-537)

QY 96 IleTyrLeuValGlyMetMetGlySerGlyLysSerThrValGlyLysIleMetSerGlu 115
DB 37 ATTTTCGATTGCTGTTTCAATTTTCACTATTACGGTGAAAAAGCTTCGAAAAAGTGGAGTCAGAT 96
QY 116 ValLeuGlyTyrSerPhePheAspSerAspLysLeuValGluGlnAlaValGlyMetPro 135
DB 97 TGTTCACAAAAGAGATGTAGATTAGATACAAAGATTGAGGAACATATCCAAATTG--- 153
QY 136 SerValAlaGlnIlePheLysValHisSerGluAlaPhePheArgAspAsnGluSerSer 155
DB 154 TCGATTGCCGAATATTTTCACTATTACGGTGAAAAAGCTTCGAAAAAGTGGAGTCAGAT 213
QY 156 ValLeuArgAspLeuSerSerMetArgArgLeuValValAlaThrGlyGlyAlaVal 175
DB 214 ATTTTAAGAAGCTGTCAAATGAAGCAAA---ATTATCGCGACTGGTGAGGGATTGTC 270

QY 176 IleArgProIleAsnTrpArgTyrMetLys---ArgGlyLeuSerValTrpLeuAspVal 194
DB 271 CAAAGTGCAAAAATCGGGTTTTAAACACACGCCGATGTTCTTTACTTAGAGGCA 330
QY 195 ProLeuAspAlaLeuAlaArgArgIleAlaLysValGlyThrAlaSerArgProLeu 213
DB 331 GAAGCGGACTGTCTAGTTGATCGTATCCAAACAAGATGAACCGTCTATTCCGACCTTTA 387

RESULT 12

US-09-583-110-886
Sequence 886, Application US/09583110
Patent No. 6699703

GENERAL INFORMATION:

APPLICANT: Lynn Doucette-Stamm et al.
TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
TITLE OF INVENTION: Pneumoniae for Diagnostics and Therapeutics
FILE REFERENCE: PATH00-07A
CURRENT APPLICATION NUMBER: US/09/583,110
CURRENT FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/107,433
PRIOR FILING DATE: 1998-06-30
PRIOR APPLICATION NUMBER: US 60/085,131
PRIOR FILING DATE: 1998-05-12
PRIOR APPLICATION NUMBER: US 60/051,553
PRIOR FILING DATE: 1997-07-02
NUMBER OF SEQ ID NOS: 5322
SEQ ID NO 886
LENGTH: 477
TYPE: DNA
ORGANISM: Streptococcus pneumoniae
US-09-583-110-886

Alignment Scores:
Pred. No.: 3.45e-10 Length: 477
Score: 179.00 Matches: 62
Percent Similarity: 50.00% Conservative: 27
Best Local Similarity: 34.83% Mismatches: 63
Query Match: 11.73% Indels: 26
DB: 4 Gaps: 8

US-10-660-226-10 (1-305) x US-09-583-110-886 (1-477)

QY 98 LeuValGlyMetMetGlySerGlyLysSerThrValGlyLysIleMetSerGluValLeu 117
DB 13 TTATTAGGGTTTATGGGGCTGGAATAATGACTATTGCAAGA-----GCCTTG 60
QY 118 GlyTyrSerPhePheAspSerAspLysLeuValGluGlnAlaValGlyMetProSerVal 137
DB 61 GACACTAATTACCTTGATATGATGCTCTGTGATGAGAGCGCTAGGTATG---TCCATT 117
QY 138 AlaGlnIlePheLysValHisSerGluAlaPhePheArgAspAsnGluSerSerValLeu 157
DB 118 GCGAATTTTTTCGTGAAAGGGAGAGAGACCTTTTCGTGAGGTAGATCAGAAGTCTTA 177
QY 158 ArgAspLeuSerSerMetArgArgLeuValAlaThrGlyGlyAlaValIleArg 177
DB 178 GCTGATTACTACAAACAGACCAA---GTCGTGTCACACTGGAGGAGGTGTTATTCT 234
QY 178 ProIleAsnTrpArgTyrMetLysArgGlyLeu---ServValTrpLeuAspValProLeu 196
DB 235 CAGAGAAATCGTGACTTACTCAAGACTAATACAGATAACATCTACCTGAAGCAGATTTT 294
QY 197 AspAlaLeuAlaArgArgIleAlaLysValGlyThrAlaSerArgProLeuLeuAspGln 216
DB 295 GAAACCCCTTACCAACGATATCGCTGATAAGGCAATATCAGCAGCCGCTTTTCTTAAT 354
QY 217 ProSerGlyAspProTyrAlaMetAlaPheSerLysLeuSerMetLeuAlaGlnArg 236
DB 355 AATAGCAGGAA-----GAACTAGTAGCTATTATTTTCAAGAAAGA 393
QY 237 GlyAspAlaTyrAlaAsnAlaAspValArgValSerLeuGluGluIleAlaCysLysGln 256

```
Db 394 CAGGCTTGCTAT-----GAGGAAGTGCTAGTCGGTT 426
Qy 257 GlyHisAspValSerLysLeuThrProThrAspIleAlaIleGluSerLeu 274
Db 427 -----TTGGATGTGACCAAGCTAAGCCAGAGGAATT---ATAGAGGAACCTA 471

RESULT 13
US-09-107-433-795
; Sequence 795, Application US/09107433
; Patent No. 6800744
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID
; SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE FOR DIAGNO
; THERAPEUTICS
; NUMBER OF SEQUENCES: 5206
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD/ROM ISO9660
; COMPUTER: <Unknown>
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: <Unknown>
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,433
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/ 085131
; FILING DATE: May 12, 1998
; APPLICATION NUMBER: 60/051553
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-011
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 795:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 480 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...480
; SEQUENCE DESCRIPTION: SEQ ID NO: 795:
US-09-107-433-795

Alignment Scores:
Pred. No.: 3 48e-10 Length: 480
Score: 179.00 Matches: 62
Percent Similarity: 50.00% Conservative: 27
Best Local Similarity: 34.83% Mismatches: 63
Query Match: 11.73% Indels: 26
DB: 4 Gaps: 8

US-10-660-226-10 (1-305) x US-09-107-433-795 (1-480)
Qy 98 LeuValGlyMetMetGlySerGlySerThrValGlyLysIleMetSerGluValLeu 117
Db 16 TTATTAGGGTTTATGGGGCTGGAATAATCGACTATTGCAAGA-----GGCTTG 63
```

```
Qy 118 GlyTyrSerPhePheAspSerAspLysLeuValGluGlnAlaValGlyMetProSerVal 137
Db 64 GACACTAATTACTCTTGTATATGATCTCTGATTGAAGCGCCTAGGTATG---TCCATT 120
Qy 138 AlaGlnIlePheLysValHisSerGluAlaPhePheArgAspAsnGluSerSerValLeu 157
Db 121 GCGAATTTTTCGCTGAAAGGGAGAGAGACCTTTTCGTAGGTAGTAATCAGAAGTCCTA 180
Qy 158 ArgAspLeuSerSerMetArgArgLeuValValAlaThrGlyGlyAlaValIleArg 177
Db 181 GCTGATTCTACTACAAACAGACCAA---GTCGTGTCACTGGAGGAGGAGTGTATTCT 237
Qy 178 ProIleAsnTrpArgTyrMetLysArgGlyLeu---SerValTrpLeuAspValProLeu 196
Db 238 CAGAGAAATCGTGACTTACTCAAGACTAATACAGATAACATCTACTCAAGACAGATT 297
Qy 197 AspAlaLeuAlaArgArgIleAlaLysValGlyThrAlaSerArgProLeuLeuAspGln 216
Db 298 GAAACCTCTACCAAGGTATCGAGCTGATAAGGACAATCAGCGACCGCTTTTCTAAAT 357
Qy 217 ProSerGlyAspProTyrAlaMetAlaPheSerLysLeuSerMetLeuAlaGlnArg 236
Db 358 AATAGCAAGGAA-----GAACTAGTAGCTATTATTTTCAAGAAAGA 396
Qy 237 GlyAspAlaTyrAlaAsnAlaAspValArgValSerLeuGluGluIleAlaCysLysGln 256
Db 397 CAGGCTTGCTAT-----GAGGAAGTGCTAGTCGGTT 429
Qy 257 GlyHisAspValSerLysLeuThrProThrAspIleAlaIleGluSerLeu 274
Db 430 -----TTGGATGTGACCAAGCTAAGCCAGAGGAATT---ATAGAGGAACCTA 474

RESULT 14
US-08-961-527-86/c
; Sequence 86, Application US/08961527
; Patent No. 6420135
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 391
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,527
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 86:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 19390 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
```


US-08-961-527-86

Alignment Scores:

Pred. No.: 1.36e-07 Length: 19390
Score: 177.00 Matches: 62
Percent Similarity: 49.44% Conservative: 26
Best Local Similarity: 34.83% Mismatches: 64
Query Match: 11.60% Indels: 26
DB: Gaps: 8

US-10-660-226-10 (1-305) x US-08-961-527-86 (1-19390)

```
QY 98 LeuValGlyMetMetGlySerGlyLysSerThrValGlyLysIleMetSerGluValLeu 117
Db 9389 TTATTAGGGTTTATGGGGCTGGAAAATCGACTATTGCAAGA-----GGCTTG 9342
QY 118 GlyTyrSerPhePheAspSerAspLysLeuValGluAlaValGlyMetProSerVal 137
Db 9341 GACCCTAATTACCTTGATGCTCTGATTGAGAAATCGCCTAGGTATG---TCCATT 9285
QY 138 AlaGlnIlePheLysValHisSerGluAlaPhePheArgAspAsnGluSerSerValLeu 157
Db 9284 GCGAATTTTTCGTCGTAAGGGAGAAACAGCCTTTCGTCAGGTAGTAATCAGAGTCCTA 9225
QY 158 ArgAspLeuSerSerMetArgArgLeuValValAlaThrGlyGlyAlaValIleArg 177
Db 9224 GCTGATTACTACAAACAGACCAAA---GTCGTGTCAACTGGAGGAGGTGTTTCT 9168
QY 178 ProIleAsnTrpArgTyrMetLysArgGlyLeu---SerValTrpLeuAspValProLeu 196
Db 9167 CAGAGAAATCGTGACTTACTCAAGACTAATACAGATAACATCTACCTGAGACGAGATTTT 9108
QY 197 AspAlaLeuAlaArgArgIleAlaLysValGlyThrAlaSerArgProLeuLeuAspGln 216
Db 9107 GAAACCTCTACCAAGCTATCGCAGCTGATAGGACAAATCAGGACCGCTTTTCTAAT 9048
QY 217 ProSerGlyAspProTyrAlaMetAlaPheSerLysLeuSerMetLeuAlaGlnArg 236
Db 9047 AATAGCAAGGA-----GAACTAGTAGCTATTATTTCCAGAGAAAGA 9009
QY 237 GlyAspAlaTyrAlaAsnAlaAspValArgValSerLeuGluIleAlaCysLysGln 256
Db 9008 CAGGCTTGTTAT-----GAGAAAGTGGCTAGTCGGGTT 8976
QY 257 GlyHisAspAspValSerLysLeuThrProThrAspIleAlaIleGluSerLeu 274
Db 8975 -----TTGATGTGACCAAGCTAGCCAGAGGAATTT---ATAGAGGAACCTA 8931
```

RESULT 15

```
US-09-134-000C-1574
; Sequence 1574, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; PRIORITY FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1574
; LENGTH: 513
; TYPE: DNA
; ORGANISM: Enterococcus faecalis
US-09-134-000C-1574
```

Alignment Scores:

Pred. No.: 6.64e-09 Length: 513
Score: 167.50 Matches: 52
Percent Similarity: 48.91% Conservative: 38

Best Local Similarity: 28.26% Mismatches: 73
Query Match: 10.98% Indels: 21
DB: Gaps: 6

US-10-660-226-10 (1-305) x US-09-134-000C-1574 (1-513)

```
QY 96 IleTyrLeuValGlyMetMetGlySerGlyLysSerThrValGlyLysIleMetSerGlu 115
Db 16 ATTGTTTAAATTGGTTTCATGGTTCGGGTAAACAACATATCGGCCAAAGTTGGCCAAT 75
QY 116 ValLeuGlyTyrSerPhePheAspSerAspLysLeuValGluAlaValGlyMetPro 135
Db 76 AAACGAGATGCTCATCTTGAATTAGATACAGCGTTAATTTGAAAAAATAGAGCGC--- 132
QY 136 SerValAlaGlnIlePheLysValHisSerGluAlaPhePheArgAspAsnGluSerSer 155
Db 133 TCATCTCTGACTATTTCGAAAAAATATGTTGAGGAGCTTTCGAGAACAGAAACCCAA 192
QY 156 ValLeuArgAspLeuSerSerMetArgArgLeuValValAlaThrGlyGlyAlaVal 175
Db 193 CTTTAAAGGAGCTGTCAAAA---AATACAGCGCTCTTCACTGGGGCGGATTTGTT 249
QY 176 IleArgProIleAsnTrpArgTyrMetLysArgGlyLeuSerVal---TrpLeuAspVal 194
Db 250 GTCCGACCAAGAAATCGTAGCTTATTAATAATCTTTTCAGCAAGTGAATTTTATCATCG 309
QY 195 ProLeuAspAlaLeuAlaArgArgIleAlaLysValGlyThrAlaSerArgProLeuLeu 214
Db 310 ACACAGAGAGCTGTTTAAAGAAATCAGAGAGATACAGAAACCAACGCGCTTAGCT 369
QY 215 AspGlnProSerGlyAspProTyrAlaMetAlaPheSerLysLeuSerMetLeuAlaGln 234
Db 370 ATAGAACGTTCTTCA-----AAAGAAATCATTACTTTGTTTGTAG 408
QY 235 GlnArgGlyAspAlaTyrAlaAsnAlaAspValArgValSerLeuGluIleAlaCys 254
Db 409 TCTCGTAAAAATTTTAT-----GAGAA-----TGT 435
QY 255 LysGlnGlyHisAspAspValSerLysLeuThrProThrAspIleAlaIleGluSerLeu 274
Db 436 GCGAAGATGACAAATTGATACGACCAATCGCTGCCGAGAGAAATTTATCAATGAAATTCG 495
QY 275 HisLysIleGlu 278
Db 496 CAACAATTAAG 507
```

Search completed: August 25, 2005, 23:30:01
Job time : 762 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 25, 2005, 20:43:32 ; Search time 721 Seconds
(without alignments)
2767.881 Million cell updates/sec

Title: US-10-660-226-10

Perfect score: 1526

Sequence: 1 MEAGGVGLALQARAAGFGSS.....ADSSASDAQESQIQRIQTL 305

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 7331713 seqs, 3271544945 residues

Total number of hits satisfying chosen parameters: 14663426

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DRV=xlh
-Q=/cgn2_1/USPTO_spool/US10660226/runat_25082005_110428_28579/app_query.fasta_1.455
-DB=Published Applications_NA -QFMT=fastcap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -LOCAL=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US10660226@cgn 1.1.480 @runat_25082005_110428_28579
-NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:*
19: /cgn2_6/ptodata/1/pubpna/US10G_PUBCOMB.seq:*
20: /cgn2_6/ptodata/1/pubpna/US10H_PUBCOMB.seq:*
21: /cgn2_6/ptodata/1/pubpna/US10I_PUBCOMB.seq:*
22: /cgn2_6/ptodata/1/pubpna/US10J_PUBCOMB.seq:*
23: /cgn2_6/ptodata/1/pubpna/US11A_PUBCOMB.seq:*
24: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
25: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
26: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1526	100.0	1059	18	US-10-425-114-30672	Sequence 30672, A
2	1526	100.0	1200	18	US-10-660-226-9	Sequence 9, Appli
3	1457	95.5	1474	20	US-10-425-115-16472	Sequence 16472, A
4	1387	90.9	1228	18	US-10-425-114-35273	Sequence 35273, A
5	1387	90.9	1241	18	US-10-425-114-31574	Sequence 31574, A
6	1387	90.9	1583	20	US-10-425-115-16474	Sequence 16474, A
7	1351.5	88.6	1273	19	US-10-767-701-13420	Sequence 13420, A
8	1301.5	85.3	1410	19	US-10-437-963-45320	Sequence 45320, A
9	1159	76.0	1061	18	US-10-660-226-27	Sequence 27, Appl
10	1046	68.5	838	20	US-10-425-115-16473	Sequence 16473, A
11	906.5	59.4	1147	19	US-10-437-963-84942	Sequence 84942, A
12	881.5	57.8	1323	18	US-10-660-226-25	Sequence 25, Appl
13	865.5	56.7	1503	20	US-10-425-115-68871	Sequence 68871, A
14	860	56.4	960	18	US-10-660-226-19	Sequence 19, Appl
15	841	55.1	899	18	US-10-660-226-11	Sequence 11, Appl
16	791.5	51.9	1223	20	US-10-739-930-4532	Sequence 4532, Ap
17	741	48.6	738	20	US-10-425-115-16469	Sequence 16469, A
18	691	45.3	1256	19	US-10-437-963-2340	Sequence 2340, Ap
19	631.5	41.4	637	21	US-10-487-901-57	Sequence 29872, A
20	625.5	41.0	1327	18	US-10-425-114-29872	Sequence 29872, A
21	625.5	41.0	1354	18	US-10-424-599-70058	Sequence 70058, A
22	620	40.6	397	19	US-10-437-963-84944	Sequence 84944, A
23	606.5	39.7	1332	20	US-10-739-930-523	Sequence 523, App
24	568	37.2	967	18	US-10-424-599-137299	Sequence 137299,
25	568	37.2	967	18	US-10-425-114-29612	Sequence 29612, A
26	553.5	36.3	544	18	US-10-660-226-15	Sequence 15, Appl
27	547	35.8	536	18	US-10-425-114-6610	Sequence 6610, Ap
28	532.5	34.9	696	19	US-10-437-963-2317	Sequence 2317, Ap
29	504	33.0	1041	19	US-10-767-701-11266	Sequence 11266, A
30	483.5	31.7	641	20	US-10-739-930-1680	Sequence 1680, Ap
31	473	31.0	1010	19	US-10-767-795-1616	Sequence 1616, Ap
32	409	26.8	492	18	US-10-424-599-25738	Sequence 25738, A
33	405	26.5	657	18	US-10-424-599-62654	Sequence 62654, A
34	385.5	25.3	660	21	US-10-487-901-5640	Sequence 5640, Ap
35	376	24.6	786	18	US-10-424-599-70060	Sequence 70060, A
36	374	24.5	309	11	US-08-864-408A-8601	Sequence 8601, Ap
37	368	24.1	329	21	US-10-487-901-5638	Sequence 5638, Ap
38	364.5	23.9	1183	18	US-10-660-226-21	Sequence 21, Appl
39	362.5	23.8	614	19	US-10-767-795-1617	Sequence 1617, Ap
40	347.5	22.8	573	19	US-10-021-323-7361	Sequence 7361, Ap
41	345	22.6	1096	18	US-10-424-599-11365	Sequence 11365, A
42	344.5	22.6	365	20	US-10-425-115-151552	Sequence 151552,
43	333	21.8	958	18	US-10-424-599-82116	Sequence 82116, A
44	319	20.9	1154	18	US-10-424-599-11361	Sequence 11361, A
45	318.5	20.9	570	17	US-10-369-493-26307	Sequence 26307, A

ALIGNMENTS

RESULT 1

US-10-425-114-30672
; Sequence 30672, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovacic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313) B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28

```
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 30672
; LENGTH: 1059
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMFLB73062D01_FLI
US-10-425-114-30672

Alignment Scores:
Pred. No.: 4,5e-167 Length: 1059
Score: 1526.00 Matches: 305
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 18 Gaps: 0

US-10-660-226-10 (1-305) x US-10-425-114-30672 (1-1059)

QY 1 MetGluAlaGlyGlyValGlyLeuAlaLeuGlnAlaArgAlaAlaGlyPheGlySerSer 20
Db 29 ATGAGCGCGGGCGCTCGCGCTGCGCGCTGCGAGCGCGGGCGCGGCTTCGGCTCCAGC 88
QY 21 ArgHisArgGlyGlyLeuGlnAlaProThrGlySerLeuArgValAlaAspProAlaGly 40
Db 89 CGGACCGGGCGGCTACAGCGGCCACCGGAGCTCGAGAGCTCGTACCGCGGGGA 148
QY 41 ProAlaValAlaValArgAlaArgGlySerLysProValAlaProLeuArgLeuArgAla 60
Db 149 CTGCGGCTCGCTGTGCGGCTCGCGGCTCAAGCCCGCTCGCACCGCTCGGACTCGCGTGG 208
QY 61 LysLysSerSerGlyGlyHisGluAsnSerHisAsnSerValAspGluAlaLeuLeuLeu 80
Db 209 AGAATATCGTCGGAGGTATGAAACTCGCAACTCGCTTGTGACCAAGCTCTCTGTTG 268
QY 81 LysArgLysSerGluGluValLeuPheTyrLeuAsnGlyArgCysIleTyrLeuValGly 100
Db 269 AAGAGAAATCAGAGAAGTCTGTCTACTTGAACGGGAGGTGATTTACCTAGTAGGA 328
QY 101 MetMetGlySerGlyLysSerThrValGlyLysIleMetSerGluValLeuGlyTyrSer 120
Db 329 ATGATGGTCTTGAAAAAGTACTGTGGGGAAGTATGCTGAAGTCTTGGGTTATTGG 388
QY 121 PhePheAspSerAspLysLeuValGluGlnAlaValGlyMetProSerValAlaGlnIle 140
Db 389 TTTCTTTGATGACAAAGTTAGTGGAGCAAGCTGTTGGAAATGCCATCAGTTGCCAAATA 448
QY 141 PheLysValHisSerGluAlaPhePheArgAspAsnGluSerSerValLeuArgAspLeu 160
Db 449 TTCAAGTCCATAGTAGAAGCTTCTTTCGGGATAATGAGAGTAGTGTCTTGAGAGATTG 508
QY 161 SerSerMetArgArgLeuValAlaThrGlyGlyGlyAlaValIleArgProIleAsn 180
Db 509 TCTCCATGCGACGATTAGTGTGTGACCGGAGGTGCTGTATTCGACCAATAAC 568
QY 181 TrpArgTyrMetLysArgGlyLeuSerValTrpLeuAspValProLeuAspAlaLeuAla 200
Db 569 TGGAGATATATGAAGAGGGCCCTATCTGTTGGTTAGATGTGCCCTTGGATGCTCTGTCT 628
QY 201 ArgArgIleAlaLysValGlyThrAlaSerArgProLeuLeuAspGlnProSerGlyAsp 220
Db 629 AGGCGTATTGCTAAAGTGGAACTGCTCTCGTCTCTTTCGACCAACCACTGGTGAT 688
QY 221 ProTyrAlaMetAlaPheSerLysLeuSerMetLeuAlaGlnGlnArgGlyAspAlaTyr 240
Db 689 CCGTAGCAATGGCCCTTTCTAAGCTCAGCATGCTTTCGACAGCAAAAGGGGTGATGCTTAT 748
QY 241 AlaAsnAlaAspValArgValSerLeuGluIleAlaCysLysGlnGlyHisAspAsp 260
Db 749 GCAATGCGAGATGAAGGGTTCTCTGGAAGAGATTGCATGTAAACAAGTATATGAT 808
QY 261 ValSerLysLeuThrProThrAspIleAlaIleGluSerLeuHisLysIleGluSerPhe 280
Db 943
```

```
809 GTCTTAAGCTGACACCTACTGATATTGCAATTGAGTCACTTATAGATCGAGAGCTTC 868
QY 281 ValIleGluHisThrAlaAspSerSerAlaSerAspAlaGlnAlaGlySerGlnIleGln 300
Db 869 GTCATCGAGCACACTGCTGATAGTTTCAGTAGCAGCGGCAAGCTGAGTGCAGATCCAG 928
QY 301 ArgIleGlnThrLeu 305
Db 929 AGGATACAGACCTTG 943

RESULT 2
US-10-660-226-9
; Sequence 9, Application US/10660226
; Publication No. US20040064848A1
; GENERAL INFORMATION:
; APPLICANT: Cahoon, Rebecca E.
; APPLICANT: Falco, S. Carl
; APPLICANT: Famodu, Layo O.
; APPLICANT: Hitz, William D.
; APPLICANT: Rendina, Alan
; TITLE OF INVENTION: Chorismate Biosynthesis Enzymes
; FILE REFERENCE: BB-1159-C
; CURRENT APPLICATION NUMBER: US/10/660,226
; PRIOR FILING DATE: 2003-09-11
; PRIOR APPLICATION NUMBER: US/09/354,501
; PRIOR FILING DATE: 1999-07-16
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/093,611
; PRIOR FILING DATE: EARLIER FILING DATE: July 21, 1998
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 9
; LENGTH: 1200
; TYPE: DNA
; ORGANISM: Zea mays
US-10-660-226-9

Alignment Scores:
Pred. No.: 5,37e-167 Length: 1200
Score: 1526.00 Matches: 305
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 18 Gaps: 0

US-10-660-226-10 (1-305) x US-10-660-226-9 (1-1200)

QY 1 MetGluAlaGlyGlyValGlyLeuAlaLeuGlnAlaArgAlaAlaGlyPheGlySerSer 20
Db 144 ATGAGAGCGGGCGGCTCGCGCTGCGCTGCGAGCGCGGGCGCGGCTTCGGCTCCAGC 203
QY 21 ArgHisArgGlyGlyLeuGlnAlaProThrGlySerLeuArgValAlaAspProAlaGly 40
Db 204 CGGACCGGGCGGCTACAGCGGCCACCGGAGCTCGTAGAGTCTGACCGCGGGA 263
QY 41 ProAlaValAlaValArgAlaArgGlySerLysProValAlaProLeuArgLeuArgAla 60
Db 264 CTGCGGCTCGCTGTGCGGCTCGCGGCTCAAGCCCGTCCGACCGCTCCGACTCCGCTGG 323
QY 61 LysLysSerSerGlyGlyHisGluAsnSerHisAsnSerValAspGluAlaLeuLeuLeu 80
Db 324 AAGAAATCGTCGGAGGTATGAAACTCGCAACTCGCTTGTGAGCAAGCTCTCTGTTG 383
QY 81 LysArgLysSerGluGluValLeuPheTyrLeuAsnGlyArgCysIleTyrLeuValGly 100
Db 384 AAGAGAAATCAGAGAAAGTCTGTCTTACTTTGAAACGGGAGGTATTTACCTAGTAGGA 443
QY 101 MetMetGlySerGlyLysSerThrValGlyLysIleMetSerGluValLeuGlyTyrSer 120
Db 444 ATGATGGGTTCTGGAAAAAGTACTGTGGGAAAGTATGCTGAAAGTCTTGGGTTATTGG 503
QY 121 PhePheAspSerAspLysLeuValGluGlnAlaValGlyMetProSerValAlaGlnIle 140
Db 504 TTTCTTTGATGACAAAGTTAGTGAGCAAGCTGTTGGAATGCCATCAGTTGCCAAATA 563
```

QY 141 PheLysValHisSerGluAlaPhePheArgAspAsnGluSerSerValLeuArgAspLeu 160
Db 564 TTCAGAGTCCATAGTGAAGCTCTCTTCGCGATTAATGAGAGTAGTGTCTTTGAGAGATTG 623

QY 161 SerSerMetArgGluValValAlaThrGlyGlyGlyAlaValIleArgProIleAsn 180
Db 624 TCTCCTCATGCGACGATAGTGTGTGCGACGGAGGTGGTGTCTGTATTCGACCAATTAAC 683

QY 181 TrpArgTyrMetLysArgGlyLeuSerValTrpLeuAspValProLeuAspAlaLeuAla 200
Db 684 TGGAGATATATGAGAGGGGCTATCTGTGTGTTAGATGTGCCCTTGGATGCTCTTGCT 743

QY 201 ArgArgIleAlaLysValGlyThrAlaSerArgProLeuLeuAspGlnProSerGlyAsp 220
Db 744 AGCGGTATTGCTAAAGTGGAACTGCCTCTCGTCTCTTCTGGACCAACCATCTCGTGAT 803

QY 221 ProTyrAlaMetAlaPheSerLysLeuSerMetLeuAlaGlnArgGlyAspAlaTyr 240
Db 804 CCGTAGCGAATGGCCCTTTCTTAAGCTCAGCATGCTTGCACAGCAAAAGGGGTGATGCTTAT 863

QY 241 AlaAsnAlaAspValArgValSerLeuGluGluIleAlaCysLysGlnGlyHisAspAsp 260
Db 864 GCAAAATGCAGATGTAAGGGTTCTCTGGAAGAGATTGCATGTAAACAAGGTCTATGATGAT 923

QY 261 ValSerLysLeuThrProThrAspIleAlaIleGluSerLeuHisLysIleGluSerPhe 280
Db 924 GTCTCTAAGCTGACACCTACTGATATTGCAATTGAGTCACTTCAATAAGATCGAGAGCTTC 983

QY 281 ValIleGluHisThrAlaAspSerSerAlaSerAspAlaGlnAlaGluSerGlnIleGln 300
Db 984 GTCATCGACACACTCTGATAGTTTCAGCTAGCGACGCGCAAGCTGAGTGCAGATCCAG 1043

QY 301 ArgIleGlnThrLeu 305
Db 1044 AGGATACAGACCTTG 1058

RESULT 3

US-10-425-115-16472
; Sequence 16472, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(5322)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 16472
; LENGTH: 1474
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_115020C.1
US-10-425-115-16472

Alignment Scores:

Pred. No.: 7,59e-159 Length: 1474
Score: 1457.00 Matches: 293
Percent Similarity: 96.72% Conservative: 2
Best Local Similarity: 96.07% Mismatches: 8
Query Match: 95.48% Indels: 2
DB: 20 Gaps: 1

US-10-660-226-10 (1-305) x US-10-425-115-16472 (1-1474)

QY 1 MetGluAlaGlyValGlyValGlyLeuAlaLeuGlnAlaArgAlaAlaGlyPheGlySerSer 20
Db 96 ATGGAGCGGGGGCGTGGCGCTGCAGACGCGGGCGGCGCTTCGGCTCCGGA 155

QY 21 ArgHisArgGlyGlyLeuGlnAlaProThrGlySerLeuArgValAlaAspProAlaGly 40
Db 156 CAGCGCGGGCGGCTCAGTCCGCCATCGGAGCGCTGAGAGTCCGTGACCCGCGGGA 215

QY 41 ProAlaValAlaValArgAlaArgGlySerLysProValAlaProLeuArgLeuArgAla 60
Db 216 CTGCGGTTCGCGTCCGGGTTCGCGGTTCGAAGCCGCTCGTACCG-----CTCCGTGG 269

QY 61 LysLysSerSerGlyGlyHisGluLysSerHisAsnSerValAspGluAlaLeuLeuLeu 80
Db 270 AAGAATCTGCGGAGGTCATGAAACTTCGTTGAGCAAGACATCTCTGTTG 329

QY 81 LysArgLysSerGluGluValLeuPheTyrLeuAsnGlyArgCysIleTyrLeuValGly 100
Db 330 AAGAGAAAATCAGAAGAAGTCTCTTCTTAAACGGGAGGTGATTTTACTTAGTGGGA 389

QY 101 MetMetGlySerGlyLysSerThrValGlyLysIleMetSerGluValLeuGlyTyrSer 120
Db 390 ATGATGGGTCTCGGAAAAGTACTGTGGGGAAGATCATGTCTGAAGTCTTGGGTATTTCG 449

QY 121 PhePheAspSerAspLysLeuValGluGlnAlaValGlyMetProSerValAlaGlnIle 140
Db 450 TTCTTTGATGTCACAAATTAGTGAGGACAGCTGTGGNATGCCATCAGTTCGCCAATA 509

QY 141 PheLysValHisSerGluAlaPhePheArgAspAsnGluSerSerValLeuArgAspLeu 160
Db 510 TTCAAGGTCCATAGTGAAGCTCTTTCGGGATATAGAGTAGTGTCTTTGAGAGATTG 569

QY 161 SerSerMetArgArgLeuValValAlaThrGlyGlyGlyAlaValIleArgProIleAsn 180
Db 570 TCTCTCATGCGACGATAGTATTGTTGCCACCGAGGTGGTGTGTTATTCGACCAATTAC 629

QY 181 TrpArgTyrMetLysArgGlyLeuSerValTrpLeuAspValProLeuAspAlaLeuAla 200
Db 630 TGGAGATATATGAGAGGGGCTATCTGTTTGGTAGATGTGCCCTTGGATGCTCTTGCT 689

QY 201 ArgArgIleAlaLysValGlyThrAlaSerArgProLeuLeuAspGlnProSerGlyAsp 220
Db 690 AGCGGTATTGCTTAAAGTGGAACTGCCTCTCGTCTCTTCTGGACCAACCATCTGCTGAT 749

QY 221 ProTyrAlaMetAlaPheSerLysLeuSerMetLeuAlaGlnGlnArgGlyAspAlaTyr 240
Db 750 CCGTAGCGCAATGGCCTTTTCTAAGCTCAGCATGCTTGCACAGCAAGGGGTGATGCTTAT 809

QY 241 AlaAsnAlaAspValArgValSerLeuGluGluIleAlaCysLysGlnGlyHisAspAsp 260
Db 810 GCAAAATGCAGATGTAAGGGTTTCTCTGGAAGAGATTGCATGTAAACAAGGTCTATGATGAT 869

QY 261 ValSerLysLeuThrProThrAspIleAlaIleGluSerLeuHisLysIleGluSerPhe 280
Db 870 GTCTCTAAGCTGACACCTACTGATATTGCAATTGAGTCACTTCAATAAGATCGAGAGCTTC 929

QY 281 ValIleGluHisThrAlaAspSerSerAlaSerAspAlaGlnAlaGluSerGlnIleGln 300
Db 930 GTCATCGACACACTGCTGATAGTTTCAGCTAGCGACGCGCAAGCTGAGTGCAGATCCAG 989

QY 301 ArgIleGlnThrLeu 305
Db 990 AGGATACAGACCTTG 1004

RESULT 4

US-10-425-114-35273
; Sequence 35273, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

; FILE REFERENCE: 38-21(53313)B

; CURRENT APPLICATION NUMBER: US/10/425,114

; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 73128

; SEQ ID NO 35273

; LENGTH: 1228

; TYPE: DNA

; ORGANISM: Zea mays subsp. mexicana

; FEATURE:

; OTHER INFORMATION: Clone ID: UC-ZMFLTEOSINTE77E09_FLI

US-10-425-114-35273

Alignment Scores:

Pred. No.: 8,09e-151 Length: 1228
Score: 1387.00 Matches: 278
Percent Similarity: 94.10% Conservative: 9
Best Local Similarity: 91.15% Mismatches: 16
Query Match: 90.89% Indels: 2
DB: 18 Gaps: 1

US-10-660-226-10 (1-305) x US-10-425-114-35273 (1-1228)

```
QY 1 MetGluAlaGlyValGlyLeuAlaLeuGlnAlaArgAlaAlaGlyPheGlySerSer 20
Db 165 ATGGAGCGGGGGCGTGGCGCTGGCGCTGCAGACGGGGCGGGCGCTTCGGCTCCGGC 224
QY 21 ArgHisArgGlyLeuGlnAlaProThrGlySerLeuArgValAlaAspProAlaGly 40
Db 225 CAGCGCGGGGGCGCTACAGTCCGCCCATTTGGGAGGCTGAGAGTCTGCTGAACCGGGGGA 284
QY 41 ProAlaValAlaValArgAlaArgGlySerLysProValAlaProLeuArgLeuArgAla 60
Db 285 GCTGCGGTTGCGGTGGGGTTCGCGGCTCCAGCCGCTCGTACCG-----CTCCGTGG 338
QY 61 LysLysSerSerGlyHisGlnAsnSerHisAsnSerValAspGluAlaLeuLeuLeu 80
Db 339 AAGAATATCCGGAGGTCATGAATACTTGCATACTCCGTGACGAACTCTCTCTGTG 398
QY 81 LysArgLysSerGluValLeuPheTyrLeuAsnGlyArgCysIleTyrLeuValGly 100
Db 399 AAGAGAAATACAGAAAGTCTGTCTACTTAAACGGGAGGTGATTTACTTAGTGGGA 458
QY 101 MetMetGlySerGlyLysSerThrValGlyLysIleMetSerGluValLeuGlyTyrSer 120
Db 459 ATGATGGGTTCTGGAAGAAAGTACTGTGGGGAAGATCATGTCTGAAGTCTTGGGTTATTCG 518
QY 121 PhePheAspSerAspLysLeuValGluGlnAlaValAlaGlyMetProSerValAlaGlnIle 140
Db 519 TTCTTTGATAGTACAAATAGTGGAGCAAGCTGTTGGATGCCCTTCAGTTGCTCAATA 578
QY 141 PheLysValHisSerGluAlaPhePheArgAspAsnGluSerSerValLeuArgAspLeu 160
Db 579 TTCAAAGTTACAGTCAAGCTCTCTTTTCGGGTAATGAGAGTAGCGCTTTGAGGATCTG 638
QY 161 SerSerMetArgArgLeuValAlaThrGlyGlyAlaValIleArgProIleAsn 180
Db 639 TCCTCCATCGACGATAGTTGTTCACCGGAGGTGGTGTGTATCCGACAGTTAAC 698
QY 181 TrpArgTyrMetLysArgGlyLeuSerValTrpLeuAspValProLeuAspAlaLeuAla 200
Db 699 TCGAATATATGAAGAGGCTATCCGTTTGGTTAGATGTGCCCTTGGATGCTCTGCT 758
QY 201 ArgArgIleAlaLysValGlyThrAlaSerArgProLeuLeuAspGlnProSerGlyAsp 220
Db 759 AGCGGTATTGCTAAAGTGGGAACCGCTTCTCGCTCTCTTCGTGACCAACCGTCGGTAT 818
QY 221 ProTyrAlaMetAlaPheSerLysLeuSerMetLeuAlaGlnGlnArgGlyAspAlaTyr 240
Db 819 CCATACACATGGCCCTTTCTAGCTCAGCATGCTTCAGAGCAAGGGGTGATGCTTAT 878
QY 241 AlaAsnAlaAspValArgValSerLeuGluIleAlaCysLysGlnGlyHisAspAsp 260
```

```
Db 879 GCAAAATCGGATGTAAAGGTTTCTCTGGAAGAGATTGTCATCTAAACAAGGTCATGCGCAT 938
QY 261 ValSerLysLeuThrProThrAspIleAlaIleGluSerLeuHisIleGluSerPhe 280
Db 939 GTCTCTAAAGCTGATCCGACTGATATCGCAATTGAGTCACCTCATTAAGATCGAGATTTC 998
QY 281 ValIleGluHisThrAlaAspSerSerAlaSerAspAlaGlnAlaGluSerGlnIleGln 300
Db 999 GTCATCGAGACGCTGCTGATATCCAGCTAGCGACTCGCAAGCTGAGTCACAGATCCAA 1058
QY 301 ArgIleGlnThrLeu 305
Db 1059 AGGATACAGACCTTG 1073
RESULT 5
US-10-425-114-31574
; Sequence 31574, Application US/10425114
; Publication NO. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 31574
; LENGTH: 1241
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMFLB73191D05_FLI
US-10-425-114-31574
```

Alignment Scores:
Pred. No.: 8,21e-151 Length: 1241
Score: 1387.00 Matches: 278
Percent Similarity: 94.10% Conservative: 9
Best Local Similarity: 91.15% Mismatches: 16
Query Match: 90.89% Indels: 2
DB: 18 Gaps: 1

US-10-660-226-10 (1-305) x US-10-425-114-31574 (1-1241)

```
QY 1 MetGluAlaGlyGlyValGlyLeuAlaLeuGlnAlaArgAlaAlaGlyPheGlySerSer 20
Db 161 ATGGAGCGGGGGCGTGGCGCTGGCGCTGCAGACGGGGCGGGCGCTTCGGCTCCGGC 220
QY 21 ArgHisArgGlyGlyLeuGlnAlaProThrGlySerLeuArgValAlaAspProAlaGly 40
Db 221 CAGCGCGGGGGCGCTACAGTCCGCCCATCGGAGGCTGAGAGTCTGCTGAACCGGGGA 280
QY 41 ProAlaValAlaValArgAlaArgGlySerLysProValAlaProLeuArgLeuArgAla 60
Db 281 GCTGCGGTTGCGGTGGGGTTCGCGGCTCCAGCCGCTCGTACCG-----CTCCGTGG 334
QY 61 LysLysSerSerGlyHisGlnAsnSerHisAsnSerValAspGluAlaLeuLeuLeu 80
Db 335 AAGAATATATCCGAGGTCATGAAAACCTGCATTAATCTCGTTGACGAAGCTCTCTCTGTG 394
QY 81 LysArgLysSerGluGluValLeuPheTyrLeuAsnGlyArgCysIleTyrLeuValGly 100
Db 395 AAGAGAAATACAGAAAGATTCGTCTTACTTTAAACGGGAGGTGATTTACTTAGTGGGA 454
QY 101 MetMetGlySerGlyLysSerThrValGlyLysIleMetSerGluValLeuGlyTyrSer 120
Db 455 ATGATGGGTTCTGGAAAAAGTACTGTGGGGAAGATCATGTCTGGAAGTCTTGGGTTATTCG 514
```

```
QY 121 PhePheAspSerAspLysLeuValGluGlnAlaValGlyMetProSerValAlaGlnIle 140
Db 515 TTCTTTGATAGTACAAATAGTGGAGCAAGCTGTGTGGAAATGCTTTCAGTTGCTCAAAATA 574
QY 141 PheLysValHisSerGluAlaPhePheArgAspAsnGluSerSerValLeuArgAspLeu 160
Db 575 TTCAAAGTTCACAGTGAAGCTTCTTTCGGGATATAGAGTAGCGTCTTGGGGATCTG 634
QY 161 SerSerMetArgArgLeuValAlaThrGlyGlyAlaValIleArgProIleAen 180
Db 635 TCCTCCATCGCAGATTAGTTGTGTCACCGGAGGTGGTGTCTATCCGACCAAGTTAAC 694
QY 181 TrpArgTyrMetLysArgGlyLeuSerValTrpLeuAspValProLeuAspAlaLeuAla 200
Db 695 TGGAAATATATGAAGTGGGACCGCTATCCGTTTGGTTAGATGTGCCCTTTGGATGCTCTTGTCT 754
QY 201 ArgArgIleAlaLysValGlyThrAlaSerArgProLeuLeuAspGlnProSerGlyAsp 220
Db 755 AGGCGGATTTGCTAAAGTGGGACCGCTTCTGTCCTTCTTGGACCAACCGTCCGGTGAT 814
QY 221 ProTyrAlaMetAlaPheSerLysLeuSerMetLeuAlaGlnGlnArgGlyAspAlaTyr 240
Db 815 CCATACACAATGCGCTTTTAAAGCTCAGCATGCTTTCGAGAGCAAAAGGGGTGATGCTTAT 874
QY 241 AlaAsnAlaAspValArgValSerLeuGluIleAlaCysLysGlnGlyHisAspAsp 260
Db 875 GCAAAATGCGGATGTAAAGGTTTCTCTGGAAGAGATTGCATCTAAACAAGGTCATGCGCAT 934
QY 261 ValSerLysLeuThrProThrAspIleAlaIleGluSerLeuHisValIleGluSerPhe 280
Db 935 GTCTCTAAGCTGATGCGGACTGATATCGCAATGAGTCACTTCATTAAGATCGAGAGTTTC 994
QY 281 ValIleGluHisThrAlaAspSerSerAlaSerAspAlaGlnAlaGluSerGlnIleGln 300
Db 995 GTCATCGAGCAGCTGCTGATTAATCCAGCTAGCGACTCGCAAGCTGAGTGCACAGATCCAA 1054
QY 301 ArgIleGlnThrLeu 305
Db 1055 AGGATACAGACCTTG 1069
```

RESULT 6

```
US-10-425-115-16474
; Sequence 16474, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants
; FILE REFERENCE: 38-21(5322)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 16474
; LENGTH: 1583
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRF4577_115022C.1
US-10-425-115-16474
```

Alignment Scores:

Pred. No.:	1,166-150	Length:	1583
Score:	1387.00	Matches:	278
Percent Similarity:	94.10%	Conservative:	9
Best Local Similarity:	91.15%	Mismatches:	16
Query Match:	90.89%	Indels:	2
DB:	20	Gaps:	1

US-10-660-226-10 (1-305) x US-10-425-115-16474 (1-1583)

```
QY 1 MetGluAlaGlyGlyValGlyLeuAlaLeuGlnAlaArgAlaAlaGlyPheGlySerSer 20
Db 222 ATGGAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 281
QY 21 ArgHisArgGlyGlyLeuGlnAlaProThrGlySerLeuArgValAlaAspProAlaGly 40
Db 282 CAGCGCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 341
QY 41 ProAlaValAlaValArgAlaArgGlySerLysProValAlaProLeuArgLeuArgAla 60
Db 342 GCTCGGGTTGCGGTGCGGGTTCGAGCCCGTCCAGCCCGTCCGTACCG-----CTCCGTGG 395
QY 61 LysLysSerSerGlyGlyHisGluAsnSerHisAsnSerValAspGluAlaLeuLeuLeu 80
Db 396 AAGAAATCATCCGGAGGTTCATGAAATTCGATAAATTCGTTTGGACGAAGCTCTCTCTGTTG 455
QY 81 LysArgLysSerGluGluValLeuPheTyrLeuAsnGlyArgCysIleTyrLeuValGly 100
Db 456 AAGAAATATCAGAAAGAGTTCTGTCTTCTTAAACGGGAGGTGATTTTACTTAGTGGGA 515
QY 101 MetMetGlySerGlyLysSerThrValGlyLysIleMetSerGluValLeuGlyTyrSer 120
Db 516 ATGATGGGTTCTGGAAAAAGTACTCTGGGGAAGATCATGCTGGAAGTCTTGGGTTATTCG 575
QY 121 PhePheAspSerAspLysLeuValGluGlnAlaValGlyMetProSerValAlaGlnIle 140
Db 576 TTCTTTGATAGTACAAATAGTGGAGCAAGCTGTGTGGAAATGCTTTCAGTTGCTCAAAATA 635
QY 141 PheLysValHisSerGluAlaPhePheArgAspAsnGluSerSerValLeuArgAspLeu 160
Db 636 TTCAAAGTTCACAGTGAAGGCTTCTTTCGGGATATAGAGTAGCGTCTTGGAGGATCTG 695
QY 161 SerSerMetArgArgLeuValAlaThrGlyGlyAlaValIleArgProIleAen 180
Db 696 TCCTCCATCGCAGCATAGTTAGTTGTTGCCACCGGAGGTGGTGTCTATCCGACCAAGTTAAC 755
QY 181 TrpArgTyrMetLysArgGlyLeuSerValTrpLeuAspValProLeuAspAlaLeuAla 200
Db 756 TGGAAATATATGAAGAGAGGGCTATCCGTTTGGTTAGATGTGCCCTTGGATGCTCTTGTCT 815
QY 201 ArgArgIleAlaLysValGlyThrAlaSerArgProLeuLeuAspGlnProSerGlyAsp 220
Db 816 AGGCGTATTTGCTAAAGTGGGAAACCGCTTCTGTCCTCTCTTGGACCAACCGTCCGGTGAT 875
QY 221 ProTyrAlaMetAlaPheSerLysLeuSerMetLeuAlaGlnGlnArgGlyAspAlaTyr 240
Db 876 CCATACACAATGCGCTTTTCTAAGCTCAGCATGCTTTCGAGAGCAAAAGGGGTGATGCTTAT 935
QY 241 AlaAsnAlaAspValArgValSerLeuGluIleAlaCysLysGlnGlyHisAspAsp 260
Db 936 GCAAAATGCGGATGTAAAGGTTTCTCTGGAAGAGATTGCATCTAAACAAGGTCATGCGCAT 995
QY 261 ValSerLysLeuThrProThrAspIleAlaIleGluSerLeuHisValIleGluSerPhe 280
Db 996 GTCTCTAAGCTGATGCGGACTGATATCGCAATTTAGTCACTTCATTAAGATCGAGAGTTTC 1055
QY 281 ValIleGluHisThrAlaAspSerSerAlaSerAspAlaGlnAlaGluSerGlnIleGln 300
Db 1056 GTCATCGAGCAGCTGCTGATTAATCCAGCTAGCGACTCGCAAGCTGAGTGCACAGATCCAA 1115
QY 301 ArgIleGlnThrLeu 305
Db 1116 AGGATACAGACCTTG 1130
```

RESULT 7

```
US-10-767-701-13420
; Sequence 13420, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
```

; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement

; FILE REFERENCE: 38-21(53535)B

; CURRENT APPLICATION NUMBER: US/10/767,701

; CURRENT FILING DATE: 2004-01-29

; NUMBER OF SEQ ID NOS: 63128

; SEQ ID NO 13420

; LENGTH: 1273

; TYPE: DNA

; ORGANISM: Sorghum bicolor

; FEATURE:

; OTHER INFORMATION: Clone ID: SORBI-28MAY03-CLUS10471_1

US-10-767-701-13420

Alignment Scores:

Pred. No.: 1,14e-146 Length: 1273
Score: 1351.50 Matches: 279
Percent Similarity: 85.67% Conservative: 8
Best Local Similarity: 83.28% Mismatches: 13
Query Match: 88.56% Indels: 35
DB: 19 Gaps: 3

US-10-660-226-10 (1-305) x US-10-767-701-13420 (1-1273)

QY 1 MetGluAlaGlyValGlyLeuAlaLeuGlnAlaArg-----13
DB 122 ATGGAGCG-----GGCCTGGCGCTGCAGACGCGCGCGCTTCGGCTCCGGA 172
QY 13 -----13
DB 173 CGCGCGCGGGCGCTACAGTCGCCCATCGGAGCCTGAGAGTCGCTGATCGCGGGA 232
QY 14 -----AlaAlaGlyPheGlySerSerArgHisArgGlyLeuGlnAlaProThr 30
DB 233 GCTGCGGTTGCCCGGGCTTCGGCTCCGACCGCGCGGGCGGCTTACAGTCGCCCATC 292
QY 31 GlySerLeuArgValAlaAspProAlaGlyProAlaValAlaValAlaArgGlySer 50
DB 293 GGGAGCCTGAGATCGCTATCCCGCGGAGTCGCGGTTGCCGTGCGGCTCGGGTCC 352
QY 51 LysProValAlaProLeuArgLeuArgAlaLysLysSerSerGlyHisGluAsnSer 70
DB 353 AAGCCCGTCGCACG-----CTCCGTGCCAAGAAATCGTCGAGGTCATGAAACTTG 406
QY 71 HisAsnSerValAspGluAlaLeuLeuLysArgLysSerGluValLeuLeuPheTyr 90
DB 407 CATAACTCCGTTGACGAAGCTCTCCGTTGAAGAGAAATCCGAAGATTCGTCTTAC 466
QY 91 LeuAsnGlyArgCysIleTyrLeuValGlyMetMetGlySerGlyLysSerThrValGly 110
DB 467 TTGAACGGGAGGTGATTATTTACTTAGTAGAATGATGGGTTCTGGAAAAAGTACAGTGGG 526
QY 111 LysIleMetSerGluValLeuGlyTyrSerPheAspSerAspLysLeuValGluGln 130
DB 527 AAGATTATGCTCAAGCTCTGGGTATTTCGTTCTTTGACAGTGCACAAATTAGTGGAGCAA 586
QY 131 AlaValGlyMetProSerValAlaGlnIlePheLysValHisSerGluAlaPhePheArg 150
DB 587 GCTGTTGGAATGCGCTTCAGTTGCTCAAAATATTCAAGTTTCATAGCGAAGCCCTTCCTCG 646
QY 151 AspAsnGluSerSerValLeuArgAspLeuSerSerMetArgArgLeuValValAlaThr 170
DB 647 GATAATGACAGTAGTGTCTTGAGAGATTGTCCTCCATCAACGATTAGTTGTGCCACC 706
QY 171 GlyGlyValAlaValIleArgProIleAsnTrpArgTyrMetLysArgGlyLeuSerVal 190
DB 707 GCGCGTGGTGGCTGTTATCCGACCACTTAACCTGGAATAATATGAAGAAGGCCCTATCTGTT 766
QY 191 TrpLeuAspValProLeuAspAlaLeuAlaArgArgIleAlaLysValGlyThrAlaSer 210
DB 767 TGGTTAGATGCGCCCTTGATGCTCTTGTCTAGCGGTATTGCTTAAGTGGGAATCGCTCT 826
QY 211 ArgProLeuLeuAspGlnProSerGlyAspProTyrAlaMetAlaPheSerLysLeuSer 230

DB 827 CGTCTCTCTCTGGACCAACCATCTGGTGATCCATACAAATCGGCCTTCTTAAGCTCAGC 886
QY 231 MetLeuAlaGlnGlnArgGlyAspAlaTyrAlaAsnAlaAspValArgValSerLeuGlu 250
DB 887 ATGCTTGACAGACCAAGGGGTGACGCTTATGCNAATGCAGATGTAAAGGTTTCTCTAGAA 946
QY 251 GluIleAlaCysLysGlnGlyHisAspValSerLysLeuThrProThrAspIleAla 270
DB 947 GAGATTGCATCTAAGCAAGGTCACCATGATGTCTCTAAGCTGCACACCCATGATATCGA 1006
QY 271 IleGluSerLeuHisLysIleGluSerPheValIleGluHisThrAlaAspSerSerAla 290
DB 1007 ATTGAGTCACTTCAATAGATCGAGAGCTTCGTGACGAGACACACTCTCTGATAATCCAGCT 1066
QY 291 SerAspAlaGlnAlaGluSerGlnIleGlnArgIleGlnThrLeu 305
DB 1067 AGCGACTCGCAAGCTGAGTCGAGATCCAAAGGATACAGACCTTG 1111
RESULT 8
US-10-437-963-45320
; Sequence 45320, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 45320
; LENGTH: 1410
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_48297C.1
US-10-437-963-45320
Alignment Scores:
Pred. No.: 8,62e-141 Length: 1410
Score: 1301.50 Matches: 268
Percent Similarity: 90.35% Conservative: 13
Best Local Similarity: 86.17% Mismatches: 21
Query Match: 85.29% Indels: 9
DB: 19 Gaps: 4
US-10-660-226-10 (1-305) x US-10-437-963-45320 (1-1410)
QY 1 MetGluAlaGlyGlyValGlyLeuAlaLeuGlnAlaArgAlaAlaGlyPhe---GlySer 19
DB 76 ATGGAGGCG---GGCGTGGGCTGCGCTGCAGTCCGGGCGCGGGGTTCCGGCGCTCC 132
QY 20 SerArgHisArgGlyGlyLeu-----GlnAlaProThrGlySerLeuArg 34
DB 133 GACCGCCCGGAGCGCTCTACGGCGCGAGGGCGCGCGATCGGAGCTTGAGG 192
QY 35 ValAlaAspProAlaGlyProAlaValAlaValArgAlaArgGlySerIlyBProValAla 54
DB 193 CTCGCTAGCGCGCGCGTGCAGAACGCCCTGTGTGGGCTCGCGGTCGAAGCGGTCGCC 252
QY 55 ProLeuArgLeuArgAlaLysLysSerSerGlyGlyHisGluAsnSerHisAsnSerVal 74
DB 253 CCG-----CTCCGTGCCAAGAAATCGTCGGAGGTCATGAACATTGCTAACTCGGTT 306
QY 75 AspGluAlaLeuLeuLysArgLysSerGluGluValLeuPheTyrLeuAsnGlyArg 94

Db 307 GATGAAGCCCTCTTGCTAAAGAGAAATCAGAAGAAAGTCTCTCTTCTATTGTAATGACGCG 366
 QY 95 CysileTyrLeuValGlyMetMetGlySerGlySerThrValGlyLysIleMetSer 114
 Db 367 TGTATTACCTAGTCTGGAATGATGGTCTCGGAAAGTACTGTGGAAAGATCATGTCT 426
 QY 115 GluValLeuGlyTyrSerPhePheAspSerAspLysLeuValGluGlnAlaValGlyMet 134
 Db 427 GAAGTTTGGGGTATTTCGTTCTTTGATAGTGAATAATGGTCGAACAAGCTGTGGGCATG 486
 QY 135 ProSerValAlaGlnIlePheLysValHisSerGluAlaPhePheArgAspAsnGluSer 154
 Db 487 CCTTCAGTCTCAAAATTTTCAAGGTTTCAAGTTCATAGTGAAGCCCTCTTTAGGGAATATGAGAT 546
 QY 155 SerValLeuArgAspLeuSerSerMetArgTLeuValValAlaThrGlyGlyAla 174
 Db 547 AGTGTCTTGAGGGATTGTCTCAATGAAGCGATTAGTTGTTGCTACTGAGGTGTGCT 606
 QY 175 ValIleArgProIleAsnTrpArgTyrMetLysArgGlyLeuSerValTrpLeuAspVal 194
 Db 607 GTTATCCGACCATGTAACCTGGAATATACATGAAGAGGGCCTATCTCTTTGGTTGGATGTG 666
 QY 195 ProLeuAspAlaLeuAlaArgArgIleAlaLysValGlyThrAlaSerArgProLeuLeu 214
 Db 667 CCCTTGAGCGCTCTTGCTAGCGGTATTGCTAAAGTGGGACTCCCTCCCTCTCTTCTA 726
 QY 215 AspGlnProSerGlyAspProTyrAlaMetAlaPheSerLysLeuSerMetLeuAlaGln 234
 Db 727 GATCAACCATCTGGTGTATCCATACACATGGCTTTTCTTAACTCAGCATGCTCGGGAG 786
 QY 235 GlnArgGlyAspAlaTyrAlaAlaAlaAlaAspValArgValSerLeuGluLysAlaCys 254
 Db 787 CAAAGGGCGCATGCTTATGCAATGCTGATGTGAGGGTTTCTCTTGAAGAGATTGCATCT 846
 QY 255 LysGlnGlyHisAspValSerLysLeuThrProThrAspIleAlaIleGluSerLeu 274
 Db 847 AAACAGGGTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 906
 QY 275 HisLysIleGluSerPheValIleGluHisThrAlaAspSerSerAlaSerAspAlaGln 294
 Db 907 CATAAGATCGAAGCTTTGTCATTGAACATACCGTTGACAAATCCGGTTGGTGACTCCAG 966
 QY 295 AlaGluSerGlnIleGlnArgIleGlnThrLeu 305
 Db 967 GCTGACTACGTCGTGACAGAGATACAGACCTTG 999

RESULT 9

US-10-660-226-27
 ; Sequence 27, Application US/10660226
 ; Publication No. US20040064848A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Cahoon, Rebecca E.
 ; APPLICANT: Falco, S. Carl
 ; APPLICANT: Farnoud, Layo O.
 ; APPLICANT: Hitz, William D.
 ; APPLICANT: Rendina, Alan
 ; TITLE OF INVENTION: Chorismate Biosynthesis Enzymes
 ; FILE REFERENCE: BB-1159-C
 ; CURRENT APPLICATION NUMBER: US/10/660,226
 ; PRIOR FILING DATE: 2003-09-11
 ; PRIOR APPLICATION NUMBER: US/09/354,501
 ; PRIOR FILING DATE: 1999-07-16
 ; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/093,611
 ; PRIOR FILING DATE: EARLIER FILING DATE: July 21, 1998
 ; NUMBER OF SEQ ID NOS: 28
 ; SOFTWARE: Microsoft Office 97
 ; SEQ ID NO 27
 ; LENGTH: 1061
 ; TYPE: DNA
 ; ORGANISM: Triticum aestivum
 US-10-660-226-27

Alignment Scores:

Pred. No.: 2,14e-124 Length: 1061
 Score: 1159.00 Matches: 235
 Percent Similarity: 91.64% Conservative: 17
 Best Local Similarity: 85.45% Mismatches: 17
 Query Match: 75.95% Indels: 6
 DB: 18 Gaps: 3
 US-10-660-226-10 (1-305) x US-10-660-226-27 (1-1061)
 QY 32 SerLeuArgValAlaAspProAlaGlyProAlaValAlaValArgAlaArgGlySerLys 51
 Db 11 AGCTTGGGTGTCAGTGAATCTGGTGGGTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCG 70
 QY 52 ProValAlaProLeuArgLeuArgAlaLysLysSerSer---GlyGlyHisGluAnSer 70
 Db 71 CCGGTGCTCCCG---CTCCGCGCGCAAGAAATCGTCTGGAGGAGGTTCATGAGAACTTG 124
 QY 71 HisAnSerValAspGluAlaLeuLeuLysArgLysSerGluValLeuLeuPheTyr 90
 Db 125 CATAACTCCGTTGACGATGCCCTCTTGTGTAAGAGAGAAATCAGAAGAGAGTTCTTTTCCAG 184
 QY 91 LeuAnGlyArgCysIleTyrLeuValGlyMetMetGlySerGlyLysSerThrValGly 110
 Db 185 TTGAACGGTCCGTGATCTACCTAGTTGGAATGATGGTTCGGGGAAGAGTACGGTGGG 244
 QY 111 LysIleMetSerGluValLeuGlyTyrSerPhePheAspSerAspLysLeuValGluGln 130
 Db 245 AAGATCTTGGCTGAAGTTTGGTTATTCTTCTTCGACAGTGATAAATTTGGTCGAACAA 304
 QY 131 AlaValGlyMetProSerValAlaGlnIlePheLysValHisSerGluAlaPhePheArg 150
 Db 305 GCTGTGGCATGCCCTTCAGTCTCAAAATTTTCAAGGTTTCATAGTGAAGCCTTTCTTCA 364
 QY 151 AspAnGluSerSerValLeuArgAspLeuSerSerMetArgGluLeuValAlaThr 170
 Db 365 GATAATGAGAGTAGTGTCTTGAGGGATTTGCTCAATCGCGGATTAGTTGTGTACT 424
 QY 171 GlyGlyGlyAlaValIleArgProIleAsnTrpArgTyrMetLysArgGlyLeuSerVal 190
 Db 425 GGAGGTGGTGTCTTATCCGACCATTAATCGAAAAATATGAAGAGGCGCTATCTGTT 484
 QY 191 TrpLeuAspValProLeuAspAlaLeuAlaArgArgIleAlaLysValGlyThrAlaSer 210
 Db 485 TGGTTGGATGTGCCCTTCGAAGCTCTTCAAGCGCGTATTGCTTAAAGTGGGAGCTGCTCG 544
 QY 211 ArgProLeuLeuAspGlnProSerGlyAspProTyrAlaMetAlaPheSerLysLeuSer 230
 Db 545 CGTCTCTTCTAGATCAACCATCCCGTGATCCATACCAATGCGCTTTTCGAAACTCAGC 604
 QY 231 MetLeuAlaGlnGlnArgGlyAspAlaTyrAlaAsnAlaAspValArgValSerLeuGlu 250
 Db 605 ATGCTCGCGGAGCAAGGGCGGATGCTTATGCAATGCTGATGTGATGATGATGATGATGAT 664
 QY 251 GluIleAlaCysLysGlnGlyHisAspValSerLysLeuThrProThrAspIleAla 270
 Db 665 GAGATCGCATCTAAGCTGGTTCATGACGACGCTCTTAAGCTGACACCGGATTGATTTGT 724
 QY 271 IleGluSerLeuHisLysIleGluSerPheValIleGluHisThrAlaAspSerAla 290
 Db 725 CTGAGTCTGCTCCCAAGATCGAGGCTTTGTGTCGAA-----GACACCGCTGTC 775
 QY 291 SerAspAlaGlnAlaGluSerGlnIleGlnArgIleGlnThrLeu 305
 Db 776 GCCGACTCACAAACCGAATCGCAATCTCAAAGGATGCATACCTTG 820

RESULT 10

US-10-425-115-16473
 ; Sequence 16473, Application US/10425115
 ; Publication No. US20040214272A1
 ; GENERAL INFORMATION:
 ; APPLICANT: La Rosa, Thomas J.
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Zhou, Yihua

```
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 16473
; LENGTH: 838
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_115021C.1
US-10-425-115-16473
```

```
Alignment Scores:
Pred. No.: 2,12e-111 Length: 838
Score: 1046.00 Matches: 213
Percent Similarity: 99.07% Conservative: 0
Best Local Similarity: 99.07% Mismatches: 2
Query Match: 68.55% Indels: 1
DB: 20 Gaps: 0
```

US-10-660-226-10 (1-305) x US-10-425-115-16473 (1-838)

```
QY 1 MetGluAlaGlyValGlyLeuAlaLeuGlnAlaArgAlaArgPheGlySerSer 20
DB 188 ATGAGCGGGGGCGCTCGCGCTGCGAGCGCGGGCGCGGCTTCGGCTCCAGC 247
QY 21 ArgHisArgGlyLeuGlnAlaProThrGlySerLeuArgValAlaAspProAlaGly 40
DB 248 CGGCACCGGGCGCTCAGAGCGCCACCGGAGCTCGAGTCCGTCACCGGGGGA 307
QY 41 ProAlaValAlaValArgAlaArgGlySerLysProValAlaProLeuArgLeuArgAla 60
DB 308 CCTCGCGTCTGTGCGGGCTCGCGGTCCAAAGCGCTCGCACCGCTCCGACTCCGTGGC 367
QY 61 LysLysSerSerGlyGlyHisGluAenSerHisAenSerValAspGluAlaLeuLeuLeu 80
DB 368 AAGAAATCTCGGAGGTATGAAACTCGCAACTCTCGTGTGACAAAGCTCTCCCTGTTG 427
QY 81 LysArgLysSerGluValLeuPheThrValLeuAenGlyArgCysIleThrLeuValGly 100
DB 428 AAGAGAAATCAGAAGAAGTCTGTCTACTTGAACGGGAGGTGATTTACCTAGTAGGA 487
QY 101 MetMetGlySerGlyLysSerThrValGlyLysIleMetSerGluValLeuGlyTyrSer 120
DB 488 ATGATCGGTCTTGAAAAAGTACTGTGGGAAAGATTATCTCTCAAGTCTTGGGTATTTCG 547
QY 121 PhePheAspSerAspLysLeuValGluGlnAlaValGlyMetProSerValAlaGlnIle 140
DB 548 TTTCTTTGATGATGACAAGTTAGTGAGAGCAAGCTGTTGGAATGCCATCAGTTGCCAAATA 607
QY 141 PheLysValHisSerGluAlaPhePheArgAspAenGluSerSerValLeuArgAspLeu 160
DB 608 TTCAAGTCCATAGTGAAGCTTCTTTCGGGATAATGAGAGTAGTGTCTTGAGAGATTTC 667
QY 161 SerSerMetArgArgLeuValAlaThrGlyGlyAlaValIleArgProIleAen 180
DB 668 TCTCCATGCGAGATTAGTTGTTGCCACCGGAAGTGGTGTCTGTTATCCGACCAATTAAC 727
QY 181 TrpArgTyrMetLysArgGlyLeuSerValTrpLeuAspValProLeuAspAlaLeuAla 200
DB 728 TGGAGATATATGAAGAGGGCCCTATCTGTGTTGGTTAGATGTGCCCTTGGATCTCTTGCT 787
QY 201 ArgArgIleAlaLysValGlyThrAlaSerArgProLeuLeuAsp 215
DB 788 AGGCGATTGCTAAAGT-GGAACTGCTCTCGTCTCTTCTCTGAC 831
```

RESULT 11

```
US-10-437-963-84942/c
; Sequence 84942, Application US/10437963
; Publication No. US2004012343A1
```

```
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 84942
; LENGTH: 1147
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_84129C.1
US-10-437-963-84942
```

Alignment Scores:

```
Pred. No.: 5,52e-95 Length: 1147
Score: 906.50 Matches: 212
Percent Similarity: 68.12% Conservative: 23
Best Local Similarity: 61.45% Mismatches: 37
Query Match: 59.40% Indels: 74
DB: 19 Gaps: 9
```

US-10-660-226-10 (1-305) x US-10-437-963-84942 (1-1147)

```
QY 1 MetGluAlaGlyGlyValGlyLeuAlaLeuGlnAlaArgAlaAla---GlyPheGlySer 19
DB 1147 ATGAGAGCT---AGAGCGGGGCTGCGGATGCAGTCGCGGGCGGTGTCGGGGTCGAGCT 1091
QY 20 -SerArgHisArgGlyGlyLeuGlnAlaPro-----Th 30
DB 1090 GGGCCCGGGTGGGGCGGGCGGGCGCGCGCTGATCCGCTGGGGAAGCGCCACCGC 1031
QY 30 rGlySerLeuArgValAlaAspProAlaGlyProAlaValAlaValAlaArgAlaArgGlySe 50
DB 1030 GCGGAGCTCTCGGGTCGGAGGCCCGCGACCGCGCGGCC----- 988
QY 50 rLysProValAlaProLeuArg---LeuArgAlaLysLysSer----- 63
DB 987 -AAGCCCTCGCCCGCTGTACTGCTCAAGCGCTC-CAGAGTTTTTTTAAATTGGATTGGG 930
QY 64 -----SerG1 65
DB 929 GGAAGTGTCTCTTCGTGGTGATCTGGAGTAAATGTGCAACTTATGTGTGGAAATGG 870
QY 65 Y-----GlyHisGluAenSerHisAenSerValAspGluAlaLeuLeuLysAr 82
DB 869 AGTTGTTCAGGGCCACGACAGCTTCATTAACCTCAGTTGATGAAGCCCTCCTGTTAAAGAG 810
QY 82 glySerGluGluValLeuPheThrLysAenGlyArgCysIleThrLeuValGlyMetMe 102
DB 809 GAAATCAGAAGAAGTTCTATTCTACTTGAACGCGCGGTGTATTATTATTAGTTGGAATGAT 750
QY 102 rGlySerGlyLysSerThrValGlyLysIleMetSerGluValLeuGlyTyrSerPhePh 122
DB 749 GGGCTCAGGAAGAGTACAGTTGCAAGATATTAGCCGGAAGTTTGGGTATTCTCTCT 690
QY 122 eAspSerAspLysLeuValGluGlnAlaValGlyMetProSerValAlaGlnIlePhePh 142
DB 689 TCACAGTGTATAAGTTGGTTGAACAAGCAGTTGGAATGCCCTCTCTGTTGCCAAATATTCAA 630
QY 142 sValHisSerGluAlaPhePheArgAspAenGluSerSerValLeuArgAspLeuSerSe 162
DB 629 GGAGCACAGTGAAGCGGTTTTTTCAGAGATAATGAGGCA----- 593
```

QY 162 rMetArgArgLeuValValAlaThrGlyGlyAlaValIleArgProIleAsnTrpAr 182
Db 592 -----ATGAA 588
QY 182 gTyrMetLysArgGlyLeuSerValTrpLeuAspValProLeuAspAlaLeuAlaArgAr 202
Db 587 ATATATGAAGAAAGGCTATCTCTGCTGGCTGGATGTCCTTTGGATGCACCTTGCAAGGGG 528
QY 202 gIleAlaValGlyThrAlaSerArgProLeuLeuAspGlnProSerGlyAspProTy 222
Db 527 TATTGCTCAAGTGGTACTGCTCTGCTCTCTTCTGGATCAGCCATCGAGTATCCATA 468
QY 222 rAlaMetAlaPheSerLysLeuSerMetLeuAlaGlnGlnArgGlyAspAlaTyrAlaAs 242
Db 467 CACAGCGGCTTCTCAAACTCAGCATGCTTGCAGAGCAAGAGGCGATGCTTATGCAAA 408
QY 242 nAlaAspValArgValSerLeuGluIleAlaCysLysGlnGlyHisAspValSe 262
Db 407 TGCTGATGCTCGAGTTTCTCTTGAAAGAAATTCAGCTAAACAAGGCGCATGACGATGCTC 348
QY 262 rLysLeuThrProThrAspIleAlaIleGluSerLeuHisLysIleGluSerPheValI1 282
Db 347 CAACTAACACCGACCGATATGCTATTGAGGCTTACTCAGATTGAGAAATTCGTCAC 288
QY 282 eGluHis-----ThrAlaAspSerSerAlaSerAspAlaGlnAlaGluSerGlnIleGlnAr 301
Db 287 CGAGCATTCACATCAATCAAGTGGCCAGTTGGGGACTTAATAGTTGATTCGCCAGAAATCGAAG 228
QY 301 gIleGlnThrLeu 305
Db 227 GACAAAGGCGTTG 215

RESULT 12

US-10-660-226-25

; Sequence 25, Application US/10660226

; Publication No. US20040064848A1

; GENERAL INFORMATION:

; APPLICANT: Cahoon, Rebecca E.

; APPLICANT: Falco, S. Carl

; APPLICANT: Farnodu, Layo O.

; APPLICANT: Hitz, William D.

; APPLICANT: Rendina, Alan

; TITLE OF INVENTION: Chorismate Biosynthesis Enzymes

; FILE REFERENCE: BB-1159-C

; CURRENT APPLICATION NUMBER: US/10/660,226

; CURRENT FILING DATE: 2003-09-11

; PRIOR APPLICATION NUMBER: US/09/354,501

; PRIOR FILING DATE: 1999-07-16

; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/093,611

; PRIOR FILING DATE: EARLIER FILING DATE: July 21, 1998

; NUMBER OF SEQ ID NOS: 28

; SOFTWARE: Microsoft Office 97

; SEQ ID NO 25

; LENGTH: 1323

; TYPE: DNA

; ORGANISM: Triticum aestivum

US-10-660-226-25

Alignment Scores:

Pred. No.:	5,45e-92	Length:	1323
Score:	881.50	Matches:	183
Percent Similarity:	78.75%	Conservative:	32
Best Local Similarity:	67.03%	Mismatches:	52
Query Match:	57.77%	Indels:	6
DB:	18	Gaps:	2

US-10-660-226-10 (1-305) x US-10-660-226-25 (1-1323)

QY 14 AlaAlaGlyPheGlySerArgHisArgGlyGlyLeuGlnAlaProThrGlySerLeu 33
Db 22 GCGCAGAGGATTCCTCCCGGCGACAGTGGCGGGTGGAGGCTGCACCAAGTCCGCGC--- 78
QY 34 ArgValAlaAspProAlaGlyProAla-ValAlaValArgAlaArgGlySerLysProVa 53

Db 79 -----GCGCGCGCGCTGGTCTCGCACCGCGCGGAGCCGAGCAGCATCCCAT 132
QY 53 1-----AlaProLeuArgLeuArgAlaLysSerSerGlyGlyHisGluAsnSerHi 71
Db 133 CCGTGGCGCCACCCCTCAAGGCCCTGTGTGCCACAAATCGGAGGTACTTGAGAAAGCCCA 192
QY 71 sAsnSerValAspGluAlaLeuLeuLysArgLysSerGluGluValLeuPheTyrLe 91
Db 193 CTATTCTGCTGATGAGGCTCTCGTACTAAAGCAAAAGCAGAGGAGCTGCTCCCTTACCT 252
QY 91 uAsnGlyArgCysIleTyrLeuValGlyMetMetGlySerGlyLysSerThrValGlyLy 111
Db 253 GAATCAGCGCTGTTTATCTAGTTGAATGATGGGTTCCGCAAAACTACAGTTGGGAA 312
QY 111 stleMetSerGluValLeuGlyTyrSerPhePheAspSerAspLysLeuValGluGlnAl 131
Db 313 GATAATAGCTGAAGTACTAGGCTATTCTTTCACAGTGAATAGCTGGTTGAGCAGTG 372
QY 131 aValGlyMetProSerValAlaGlnIlePheLysValHisSerGluAlaPhePheArgAs 151
Db 373 TGTTGGCATACCGTGGTGGCTGAGATTTTTCAGGTCCACAGTGAAGCATTCTTCAGAGA 432
QY 151 pAsnGluSerSerValLeuArgAspLeuSerSerMetArgArgLeuValValAlaThrGl 171
Db 433 TAACGAGAGTGAAGTACTAAGGGATTGTCTGTCATGCAATGCACCGATTAAATTGTTGCAACGG 492
QY 171 yGlyGlyAlaValIleArgProIleAsnTrpArgTyrMetLysArgGlyLeuSerValTr 191
Db 493 AGGTGTCGGTGATACGACCAATCAATTTGGAGTTATATGAAGAAGGACTCATTATTG 552
QY 191 pLeuAspValProLeuAspAlaLeuAlaArgIleAlaLysValGlyThrAlaSerAr 211
Db 553 GTTAGATGTTCCATTGGACGCCCTTGCAGAGAGGATTCGTCGGTTGGTACTCGCTCAGC 612
QY 211 gProLeuLeuAspGlnProSerGlyAspProTyrAlaMetAlaPheSerLysLeuSerMe 231
Db 613 ACCCTCTCTGCATCAGGAATCTGTGTATGCTTATGCAAGGCTATGCCAAACTTACAGC 672
QY 231 tLeuAlaGlnGlnArgGlyAspAlaTyrAlaAsnAlaAspValArgValSerLeuGluGl 251
Db 673 ACTTTTGAACAAAGAAATGGATTGCTAATGCTATGATGCCGAGTTTCCCTTGAATA 732
QY 251 uIleAlaCysLysGlnGlyHisAspValSerLysLeuThrProThrAspIleAlaI1 271
Db 733 TATTGCATTCAACAAAGGACATAATGATGTGAATGCTACTTACCAAGTGCCTATCGCTAT 792
QY 271 eGluSerLeuHisLysIleGluSerPheValIleGlu 283
Db 793 TGAGGCATTGCTAAAGATGGAGGCTTCTTACTAGAG 829

RESULT 13

US-10-425-115-68871/C

; Sequence 68871, Application US/10425115

; Publication No. US20040214272A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.

; APPLICANT: Kovalic, David K.

; APPLICANT: Zhou, Yihua

; APPLICANT: Cao, Yongwei

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

; FILE REFERENCE: 38-21(53222)B

; CURRENT APPLICATION NUMBER: US/10/425,115

; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 369326

; SEQ ID NO 68871

; LENGTH: 1503

; TYPE: DNA

; ORGANISM: Zea mays

; FEATURE:

; OTHER INFORMATION: Clone ID: MRT4577_162804C.1

US-10-425-115-68871

```
Alignment Scores:
Pred. No.:      4,74e-90      Length:      1503
Score:          865.50      Matches:      180
Percent Similarity: 77.01%      Conservative: 31
Best Local Similarity: 65.69%      Mismatches: 39
Query Match:      56.72%      Indels:      24
DB:              20          Gaps:         2

US-10-660-226-10 (1-305) x US-10-425-115-68871 (1-1503)
QY 18 GlySerSerArgHisArgGlyGlyLeuGlnAlaProThrGlySer-Leu----- 33
Db 1132 GCGAGGCTCGCGTTCGGCGGACAGCGCGGAGGCTGTGCTGCTGGCGCGATAC 1073
QY 34 ----ArgValAlaAspProAlaGlyProAlaValAlaValArgAlaArgGlySerLysPr 52
Db 1072 GCGAGGCGCGGATCTCTGCC----- 1051
QY 52 oValAlaProLeuArgLeuArgAlaLysLysSerSerGlyGlyHisGluAsnSerHisAs 72
Db 1050 -----CTCGGTGGCGCGCGAATCTGCAGGAACAGAAAGGTCCACTA 1007
QY 72 nSerValAspGluAlaLeuLeuLeuLysArgLysSerGluGluValLeuPheTyrLeuAs 92
Db 1006 CTCCTGCTGATGACGCTCATCTACTACAGCAAAAGCCAGGATGTTCTGCCCTTACTTGA 947
QY 92 nGlyArgCysIleTyrLeuValGlyMetMetGlySerGlyLysSerThrValGlyLysIl 112
Db 946 TGGCCGCTTCGTTATCTTGTGGAATGATGGGTTTCAGCAAAACTACAGTTGGGAAGAT 887
QY 112 eMetSerGluValLeuGlyTyrSerPhePheAspSerAspLysLeuValGluGlnAlaVa 132
Db 886 ACTATCCGAAGTGTAGGTATTCTGTTCTTCACAGTATAAGTTGGTAGAAGGCTGT 827
QY 132 lGlyMetProSerValAlaGlnIlePheLysValHisSerGluAlaPhePheArgAspAs 152
Db 826 TGGTATTTCATCTGTGTGCTGAGATCTTTCAGCTCCATCGCAAAACATCTTCAGAGATAA 767
QY 152 nGluSerSerValLeuArgAspLysSerSerMetArgArgLeuValValAlaThrGlyGl 172
Db 766 TGAGAGTGAGGTCCTGAGGGATCTGTCAATATGCATCGGTTGGTTTGGCAACGGAGG 707
QY 172 yGlyAlaValIleArgProIleAsnTrpArgTyrMetLysArgGlyLeuSerValTrpLe 192
Db 706 TGGTGAGTGATCCGACCAATCAATTGGAGTTACATGAAGAAGGCTGACCCGTATGGTT 647
QY 192 uAspValProLeuAspAlaLeuAlaArgArgIleAlaLysValGlyThrAlaSerArgPr 212
Db 646 AGATGTCCTCCACTGGATGCATTCGAAGAAGATCGCTGCTGTAGGAACCGCGTCTCGACC 587
QY 212 oLeuLeuAspGlnProSerGlyAspProTyrAlaMetAlaPheSerLysLeuSerMetLue 232
Db 586 ACTCTTGCAATCAGGAATCTGGTGATCCTTATGCAAAAGGCTTATGCAAAACTTACATCACT 527
QY 232 uAlaGlnIleArgGlyAspAlaTyrAlaAsnAlaAspValArgValSerLeuGluIle 252
Db 526 TTTTGACAAAGAATGGACTCGTATGCTAATGCTGATGCCAGATTCTACATTGAACATAT 467
QY 252 eAlaCysLysGlnGlyHisAspAspValSerLysLeuThrProThrAspIleAlaIleGl 272
Db 466 TGCATTAAACAAGGCCAATAATGATGCTACTATCTTACACCTAGTACCATGCCATTGA 407
QY 272 uSerLeuHisLysIleGluSerPheValIleGluHisThr 285
Db 406 GGCATTGCTAAAGATGGAAGTTTCTTACCAGAGAGACC 367
```

RESULT 14

```
US-10-660-226-19
; Sequence 19, Application US/10660226
; Publication No. US2004006484A1
; GENERAL INFORMATION:
; APPLICANT: Cahoon, Rebecca E.
```

```
; APPLICANT: Falco, S. Carl
; APPLICANT: Famodu, Layo O.
; APPLICANT: Hitz, William D.
; APPLICANT: Rendina, Alan
; TITLE OF INVENTION: Chorismate Biosynthesis Enzymes
; FILE REFERENCE: BB-1159-C
; CURRENT APPLICATION NUMBER: US/10/660,226
; CURRENT FILING DATE: 2003-09-11
; PRIOR APPLICATION NUMBER: US/09/354,501
; PRIOR FILING DATE: 1999-07-16
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/093,611
; PRIOR FILING DATE: EARLIER FILING DATE: July 21, 1998
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 19
; LENGTH: 960
; TYPE: DNA
; ORGANISM: Sorghum
US-10-660-226-19

Alignment Scores:
Pred. No.:      1.1e-89      Length:      960
Score:          860.00      Matches:      174
Percent Similarity: 83.40%      Conservative: 32
Best Local Similarity: 70.45%      Mismatches: 33
Query Match:      56.36%      Indels:       8
DB:              18          Gaps:         2

US-10-660-226-10 (1-305) x US-10-660-226-19 (1-960)
QY 39 AlaGlyProAlaValAlaValArgGlySerLysProValAlaProLeuArgLeu 58
Db 8 GCGGGTCTGCGCTC-----CGTCCC---GCAAGCTGAGAGTT 43
QY 59 ArgAlaLysLysSerSerGlyGlyHisGluAsnSerHisAsnSerValAspGluAlaLeu 78
Db 44 TCGTCTCCGCAAAATCGGACGACAGCAAGAAAGTCCACTATTCTTACTGACGAGCTCTC 103
QY 79 LeuLeuLysArgLysSerGluGluValLeuPheTyrLeuAsnGlyValArgCysIleTyrLeu 98
Db 104 ATACTACAGCAAAAGGCCAGGATGTTCTCCCTTACTTGGATGGCCGATGCGTTATCTT 163
QY 99 ValGlyMetMetGlySerGlyLysSerThrValGlyLysIleMetSerGluValLeuGly 118
Db 164 GTTGAATGATGGGTTTCAGGCAAACTACAGTTGGAGATATATTAGCCGAGATTAGGT 223
QY 119 TyrSerPhePheAspSerAspLysLeuValGluGlnAlaValGlyMetProSerValAla 138
Db 224 TATTGCTTCTTTGACAGTGATAAGCTGGTAGAAGGCTGTTGGTATCTCATCTCTGCT 283
QY 139 GlnIlePheLysValHisSerGluAlaPhePheArgAspAsnGluSerSerValLeuArg 158
Db 284 GAGATCTTTCAGCTCCATAGTGAAGCATCTTCAGAGATAATAGAGTAGAGTCTCTGAGG 343
QY 159 AspLeuSerSerMetArgArgLeuValValAlaThrGlyGlyGlyValAlaValIleArgPro 178
Db 344 GATCTGTCATCAATGATCGTGGTTGTTGTGCAACCGGAGGTGGTGAGTGATCGGACCA 403
QY 179 IleAsnTrpArgTyrMetLysArgGlyLeuSerValTrpLeuAspValProLeuAspAla 198
Db 404 ATCAATTGGAGTTACATGAAGAAGGCTGACTGTGTGTTAGAGCTTCCATCGGATGCA 463
QY 199 LeuAlaArgArgIleAlaLysValGlyThrAlaSerArgProLeuLeuAspGlnProSer 218
Db 464 CTTGCAAGAAGAATTTGCTGCTGAGAACCGCATCTTCGACCACTCTTTCATCGGAATCT 523
QY 219 GlyAspProTyrAlaMetAlaPheSerLysLeuSerMetLeuAlaGlnGlnArgGlyAsp 238
Db 524 GGTGACCTTATGCAAAAGGCTTATCGGAAACTTATCATCATCTTTTGGAGCAAGAATGGAC 583
QY 239 AlaTyrAlaAsnAlaAspValArgValSerLeuGluGluIleAlaCysLysGlnGlyHis 258
Db 584 TCGTATGCTAATGCTGATGCCAGAGTTTCACTTGAACATATTGTCATTAAACAAGGCCAT 643
```

Qy 259 AspAspValSerLysLeuThrProThrAspAlaAlaIleGluSerLeuHisLysIleGlu 278
Db 644 AATGATGTCACTATCTTACACCTAGTGCCTAGCCATTGAGGCAATGCTAAAGATGGAA 703
Qy 279 SerPheValIleGluHisThr 285
Db 704 AGTTTTCTTACCGAAGACC 724

RESULT 15

US-10-660-226-11
; Sequence 11, Application US/10660226
; Publication No. US20040064848A1
; GENERAL INFORMATION:
; APPLICANT: Cahoon, Rebecca E.
; APPLICANT: Falco, S. Carl
; APPLICANT: Famodu, Layo O.
; APPLICANT: Hitz, William D.
; APPLICANT: Rendina, Alan
; TITLE OF INVENTION: Chorismate Biosynthesis Enzymes
; FILE REFERENCE: BB-1159-C
; CURRENT APPLICATION NUMBER: US/10/660,226
; CURRENT FILING DATE: 2003-09-11
; PRIOR APPLICATION NUMBER: US/09/354,501
; PRIOR FILING DATE: 1999-07-16
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/093,611
; PRIOR FILING DATE: EARLIER FILING DATE: July 21, 1998
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 11
; LENGTH: 899
; TYPE: DNA
; ORGANISM: Zea mays
US-10-660-226-11

Alignment Scores:
Pred. No.: 1,62e-87 Length: 899
Score: 841.00 Matches: 165
Percent Similarity: 87.50% Conservative: 31
Best Local Similarity: 73.66% Mismatches: 28
Query Match: 55.11% Indels: 0
DB: 18 Gaps: 0

US-10-660-226-10 (1-305) x US-10-660-226-11 (1-899)

Qy 62 LysSerSerGlyGlyHisGluAenSerHisAenSerValAspGluAlaLeuLeuLys 81
Db 11 CAATCTGCAGTGGGAACAGAAAGGTCCACTACTCTGCTGATGACGCTCTCATCTACAG 70
Qy 82 ArgLysSerGluGluValLeuPheTyrLeuAenGlyArgCysIleTyrLeuValGlyMet 101
Db 71 CAAAAGCCAGGATGTTCTGCCTTACTTGGATGGCGGTTGCTTTATCTTGTGGATG 130
Qy 102 MetGlySerGlyLysSerThrValGlyLysIleMetSerGluValLeuGlyTyrSerPhe 121
Db 131 ATGGGTTTCAGGCAAACTACAGTTGGGAAGATACTATCCGAAGTGTAGTTATTTCGTTTC 190
Qy 122 PheAspSerAspLysLeuValGluGlnAlaValGlyMetProSerValAlaGlnIlePhe 141
Db 191 TTCGACAGTGATAGTTGGTAGAAGGCTGTGGTAATTTTCATCTGTTGCTGAGATCTTT 250
Qy 142 LysValHisSerGluAlaPhePheArgAspAenGluSerSerValLeuArgAspLeuSer 161
Db 251 CAGTCCATAGCGAAACATTTCTCAGAGATATGAGAGTGAGGTCTCTGACGGATCTGTCA 310
Qy 162 SerMetArgLeuValValAlaThrGlyGlyAlaValIleArgProIleAenTrp 181
Db 311 TCAATGCATCGGTGTTGTTTTCACCGAGGTGGTGCAGTGATCCGACCAATCAATTGG 370
Qy 182 ArgTyrMetLysArgGlyLeuSerValTrpLeuAspValProLeuAspAlaLeuAlaArg 201
Db 371 AGTTACATGAAGAAAGGCTGACCGGTATGGTTAGATGTCCCACTGGATGCCTTGCAGA 430

Qy 202 ArgIleAlaLysValGlyThrAlaSerArgProLeuLeuAspGlnProSerGlyAspPro 221
Db 431 AGAATCGCTGCTGTAGGAACCGCGTCTCGACCACTTTGCAATCAGGAATCCGGTGATCCT 490
Qy 222 TyrAlaMetAlaPheSerLysLeuSerMetLeuAlaGlnGlnArgGlyAspAlaTyrAla 241
Db 491 TATGCAAGGCTTATGCAAACTTACGTCACCTTTTGGACAAAGAATGGACTCGTATGCT 550
Qy 242 AenAlaAspValArgValSerLeuGluGlnIleAlaCysLysGlnGlyHisAspAspVal 261
Db 551 AATGCTGATGCCAGAGTTTTCATTGAAACATATTGCAATAAACAAGGCCATTAATGATGC 610
Qy 262 SerLysLeuThrProThrAspIleAlaIleGluSerLeuHisLysIleGluSerPheVal 281
Db 611 ACTATACTTACACCTAGTACCATCGCCATTGAGGCATTGCTAAAGATGGAAGTTTCTT 670
Qy 282 IleGluHisThr 285
Db 671 ACCGAGAAGACC 682

Search completed: August 25, 2005, 23:32:54
Job time : 725 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 25, 2005, 20:22:28 ; Search time 3683 Seconds
(without alignments)
3152.213 Million cell updates/sec

Title: US-10-660-226-10
Perfect score: 1526
Sequence: 1 MEAGVGGLALQARAAGFGSS.....ADSSASDAQAESQIQRIQTL 305

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DRV=xlh
-Q/cgn2_1/USPTO_spool/US10660226/runat_25082005_110427_28535/app_query.fasta_1.455
-DB=EST -Qfmt=fastap -SUFFIX=rst -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10660226 @CNC 1.1 3437 @runat_25082005_110427_28535 -NCPU=6 -ICPU=3
-NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST: *
1: gb_est1: *
2: gb_est2: *
3: gb_hic: *
4: gb_est3: *
5: gb_est4: *
6: gb_est5: *
7: gb_est6: *
8: gb_gse1: *
9: gb_gse2: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Query	ID	Description
1	1346	88.2	1305	3	AY104682 Zea mays
2	1127.5	73.9	825	6	CB684025 OSJNEfl13C
3	1091.5	71.5	833	6	CB657302 OSJNEc121r
4	1050	68.8	667	6	CA078453 SCR1AM100
5	1043.5	68.4	809	6	CB656619 OSJNEc11D
6	1024.5	67.1	689	4	BJ248029 BJ248029
7	1013	66.4	804	7	CN147205 WOUND1_48
8	1012.5	66.3	797	6	CB641679 OSJNEB01H
9	1004	65.8	818	6	CB684026 OSJNEfl13C

10	990.5	64.9	772	6	CB656558
11	972	63.7	854	6	CB657303
12	968	63.4	640	4	BJ472228
13	964	63.2	672	6	CA146167
14	960	62.9	609	2	AM671996
15	957	62.7	789	6	CB656559
16	953	62.5	789	6	CB656620
17	934	61.2	619	6	CA452817
18	928.5	60.8	914	4	BG301287
19	925	60.6	620	6	CF033002
20	915	60.0	814	7	CF636791
21	909	59.6	648	5	BU986135
22	899	58.9	650	6	CA105686
23	896.5	58.7	701	7	CN144633
24	891	58.4	649	1	AV909303
25	879	57.6	733	5	CA452716
26	863.5	56.6	752	5	BQ295318
27	850	55.7	582	4	EG355018
28	848	55.6	659	6	CA112411
29	846.5	55.5	775	6	CA182703
30	843	55.2	642	6	CD670151
31	839	55.0	728	7	CN144547
32	836	54.8	612	6	CA118037
33	834	54.7	786	7	CR827364
34	832	54.5	549	7	CO524949
35	830	54.4	601	1	A1795555
36	826	54.1	776	6	CA076988
37	825	54.1	559	5	B0763798
38	822	53.9	669	6	CA248324
39	818	53.6	1066	7	CR286591
40	812.5	53.2	645	4	BG415114
41	811	53.1	729	6	CD900974
42	809	53.0	811	7	CN138158
43	805	52.8	692	4	BM736223
44	802	52.6	513	2	AM671997
45	785	51.4	838	6	CB666891

ALIGNMENTS

RESULT 1
AY104682

LOCUS Zea mays PCO108968 mRNA sequence. 1305 bp linear HTC 16-OCT-2002

DEFINITION Zea mays PCO108968 mRNA sequence.

ACCESSION AY104682

VERSION AY104682.1 GI:21207760

KEYWORDS HTC.

SOURCE Zea mays

ORGANISM Zea mays

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Direct Submission
Submitted (25-APR-2002) Maize Mapping Project, University of Missouri, Columbia, MO 65211, USA

If you are interested in getting corresponding physical clones, these are publicly available from ZmDB, www.zmdb.iastate.edu; TIGR, the searching at MSU, maizemap.org; ZmDB, www.zmdb.iastate.edu; TIGR, the maize cDNA sequences is either Virginia Walbot, Stanford or Pat Schnable, Iowa State, then clones may be requested from ZmDB: www.zmdb.iastate.edu.

FEATURES

source

1. .1305

/organism="Zea mays"

```

/mol_type="mRNA"
/db_xref="taxon:4577"
/db_xref="taxon:4577"
/clone_lib="Maize Mapping Project/DuPont Cornsensus
Library"
/note="this sequence is part of a project of EST
assemblies resulting from the application of public
contigs to seed Dupont contigs; this resource was
assembled by Dupont as part of a collaboration for the
overgo addressing of BACs in conjunction with the Maize
Mapping Project"

```

ORIGIN

```

Alignment Scores:
Pred. No.: 7.38e-130 Length: 1305
Score: 1346.00 Matches: 275
Percent Similarity: 93.11% Conservative: 9
Best Local Similarity: 90.16% Mismatches: 19
Query Match: 88.20% Indels: 3
DB: 3 Gaps: 1

US-10-660-226-10 (1-305) x AY104682 (1-1305)

QY 1 MetGluAlaGlyValGlyLeuAlaLeuAlaAlaAtgAlaAlaGlyPheGlySerSer 20
DB 197 ATGGAGCGGGGGCGCTGCGCGCTGCGAGCGCGCGCGCGCGCTTCGGCTCCGGC 256
QY 21 ArgHisArgGlyLeuGlnAlaProThrGlySerLeuArgValAlaAlaAspProAlaGly 40
DB 257 CAGCGCGGGGGCGCTACAGTCGCCCATCGGAGGCTGAGAGTCGCTGAACCGCGGGA 316
QY 41 ProAlaValAlaValAlaArgGlySerLysProValAlaProLeuArgLeuArgAla 60
DB 317 GCTGCGGGTTGCGGGGTTGCGGGGTCAGACCGCGCTGCTACCG-----CTCCGTCGG 370
QY 61 LysLysSerSerGlyGlyHisGluAenSerHisAenSerValAspGluAlaLeuLeuLeu 80
DB 371 AGAAATCATCCGAGGCTCATGAAATCTGCAATACCTCGTTGACGAGCTCTCTGTG 430
QY 81 LysArgLysSerGluGluValLeuPheTyrLeuAenGlyArgCysIleTyrLeuValGly 100
DB 431 AAGAGAAATCAGAAAGTCTGTCTTCTTAAACCGGAGGTGATTTACTTAGTGGA 490
QY 101 MetMetGlySerGlyLysSerThrValGlyLysIleMetSerGluValLeuGlyTyrSer 120
DB 491 ATGATGGTTCGTGAAAAAGTACTGTGGGGAAGATCATCTGCAAGTCTTGGGTATTTCG 550
QY 121 PhePheAspSerAspLysLeuValGluGlnAlaValGlyMetProSerValAlaGlnIle 140
DB 551 TTCTTTGATAGTACAAATATGAGGAGCAAGCTGTGGAAATGCTTCAGTTGCTCAATA 610
QY 141 PheLysValHisSerGluAlaPhePheArgAspAenGluSerSerValLeuArgAspLeu 160
DB 611 TTCAAAGTTTCACAGTGAAGCTTCTTTTCGGGATAATGAGAGTAGCGTCTTTGAGGATCTG 670
QY 161 SerSerMetArgArgLeuValAlaThrGlyGlyAlaValIleArgProIleAen 180
DB 671 TCTTCCATCGACGATAGTGTGTGTCACCGGAGT-TGNCCTTGTCTATCCGACGAGTTAC 729
QY 181 TTPATGTYrMetLysArgGlyLeuSerValTTPLeuAspValProLeuAspAlaLeuAla 200
DB 730 TGGAAATATATGAGAGGCGCTATCCGTTGTTAGATGTGCCCTTGGATGCTCTTGCT 789
QY 201 ArgArgIleAlaLysValGlyThrAlaSerArgProLeuLeuAspGlnProSerGlyAsp 220
DB 790 AGCGCATTTGCTAAAGTGGGAACCGCTTCTCGTCCTCTTTCGACCAACCGTCGGTGAT 849
QY 221 ProTyrAlaMetAlaPheSerLysLeuSerMetLeuAlaGlnArgGlyAspAlaTyr 240
DB 850 CCATACACAAATGGCCCTTTTCTAAGCTCAGCATGCTTGCAGAGCAAGGGGTGATGCTTAT 909
QY 241 AlaAenAlaAspValArgValSerLeuGluGluIleAlaCysLysGlnGlyHisAspAsp 260

```

```

Db 910 GCAATGCGGATGTAAGGGTTTCTCTGGAAGAGATGTCATTAACAAGGTCATGCGCAT 969
QY 261 ValSerLysLeuThrProThrAspIleAlaIleGluSerLeuHisLysIleGluSerPhe 280
DB 970 GTCTCTAAGCTGATGCCGACTGATATCGCAATTCAGTCACTTCAATAAGATCGAGATTTC 1029
QY 281 VallieGluHisThrAlaAspSerSerAlaSerAspAlaGlnAlaGluSerGlnIleGln 300
DB 1030 GTCATCGAGCAGCGTGTGATAATCCAGCTAGCGACTGCAAGCTGAGTGCACAGATCCA 1089
QY 301 ArgIleGlnThrLeu 305
DB 1090 AGGATACAGACCTTG 1104

RESULT 2
CB684025 825 bp mRNA linear EST 09-APR-2003
OSJNEF13C12.f OSJNEF Oryza sativa (japonica cultivar-group) cDNA
clone OSJNEF13C12 5', mRNA sequence.
ACCESSION CB684025
VERSION CB684025.1 GI:29687750
KEYWORDS EST.
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehretaceae; Oryzae; Oryza.
REFERENCE 1 (bases 1 to 825)
AUTHORS Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E.,
Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.
TITLE Large-scale identification of ESTs involved in the interaction
between rice and Magnaporthe grisea
JOURNAL Unpublished (2003)
COMMENT Contact: Rod Wing
Arizona Genomics Institute
University of Arizona
Biological Sciences West, 448A, P.O. Box 2-0088, Tucson, AZ
85721-0088, USA
Tel: 520 626 3967
Fax: 520 621 9288
Email: http://genome.arizona.edu
PCR Primers
FORWARD: gta aaa cga cgg cca gtc
BACKWARD: gga aac agc tat gac cat g
Plate: 13 row: C column: 12
Seq primer: gta aaa cga cgg cca gtc.
FEATURES
Location/Qualifiers
source
1..825
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="OSJNEF13C12"
/tissue_type="Leaf"
/dev_stage="3 week"
/lab_host="DH10B"
/clone_lib="OSJNEF"
/note="Vector: pBluescript II KS +; Site_1: EcoRI; Site_2:
XhoI; Uninfected Control"

Alignment Scores:
Pred. No.: 2.8e-107 Length: 825
Score: 1127.50 Matches: 234
Percent Similarity: 90.64% Conservative: 8
Best Local Similarity: 87.64% Mismatches: 16
Query Match: 73.89% Indels: 9
DB: 6 Gaps: 4

US-10-660-226-10 (1-305) x CB684025 (1-825)

QY 1 MetGluAlaGlyGlyValGlyLeuAlaLeuGlnAlaArgAlaAlaGlyPhe---GlySer 19

```


Db 34 ATGAGGCG---GGCGTGGGCTGGCGCTGCAGTCGCGGGCGGGGCTTCGGCGGCTCC 90
 Qy 20 SerArgHisArgGlyGlyLeu-----GlnAlaProThrGlySerLeuArg 34
 Db 91 GACCGCGCGGAGCGGCTCTACGGCGGCGGCGGCGGCGGATCGGGAGCTTGAGG 150
 Qy 35 ValAlaAspProAlaGlyProAlaValAlaValAlaArgAlaArgGlySerLysProValAla 54
 Db 151 GTCGCTGAGCGCGGCTGGCGGCGGCTGTGTGGGCTCGGGGCTCAAGCGGCTCGCC 210
 Qy 55 ProLeuArgLeuArgAlaLysLysSerSerGlyGlyHisGluAsnSerHisAsnSerVal 74
 Db 211 CCG-----CTCCGTCGCGGAGAAATCGTCGGAGGTCATGAACATGTGCACTACGTT 264
 Qy 75 AspGluAlaLeuLeuLysArgLysSerGluGluValLeuPheThrLeuAsnGlyArg 94
 Db 265 GATGAAGCCCTCTGCTAAAGAGAAATCAGAAGAGTTCTCTTCTATTGATGACGG 324
 Qy 95 CysIleThrLeuValGlyMetMetGlySerGlyLysSerThrValGlyLysIleMetSer 114
 Db 325 TGTATTACCTAGTTGGAATGATGGGTTCTGGAAAAAGTACTGTGGAAAAAGATCATGTCT 384
 Qy 115 GluValLeuGlyThrSerPhePheAspSerAspLysLeuValGluGlnAlaValGlyMet 134
 Db 385 GAAGTTTGGGTTATTCGTTCTTTGATAGTAAATTTGGTGAACAGCTGTGGCGATG 444
 Qy 135 ProSerValAlaGlnIlePheLysValHisSerGluAlaPhePheArgAspAsnGluSer 154
 Db 445 CCTTCAGTCGCTCAAAATTTCAAGGTTTCATAGTGAAGCTCTTTAGGATATAGAGT 504
 Qy 155 SerValLeuArgAspLeuSerMetArgArgLeuValValAlaThrGlyGlyAla 174
 Db 505 AGTGCTCTGAGGAGATTGCTCCTCAATGAAGCGATTAGTTGTTGCTACTCGAGGTGTGCT 564
 Qy 175 ValIleArgProIleAsnTrpArgThrMetLysArgGlyLeuSerValTrpLeuAspVal 194
 Db 565 GTTATCCGACCAATTTAACTGGAAATACATGAAGAGGGCTATCTGTTGGTTGGATGG 624
 Qy 195 ProLeuAspAlaLeuAlaArgArgIleAlaLysValGlyThrAlaSerArgProLeuLeu 214
 Db 625 CCTTGGAGCGCTTCTGCTAGGCGTATGCTTANAGTGGGACGCTCCCTCCCTCTCTTA 684
 Qy 215 AspGlnProSerGlyAspProThrAlaMetAlaPheSerLysLeuSerMetLeuAlaGln 234
 Db 685 GATCAACCATCTGCTGATCCATACACATGGCTTTTCTTAACTCAGCATGCTCGCGGAG 744
 Qy 235 GlnArgGlyAspAlaThrAlaAsnAlaAspValArgValSerLeuGluIleAlaCys 254
 Db 745 CAAAGGGCGGATGCTTATGCAATGCTGATGTGAGGGTTTCTCTTGAAGAGATTGCATCT 804
 Qy 255 LysGlnGlyHisAspAspVal 261
 Db 805 AAACAGGGTCATGATGTA 825

RESULT 3

CB657302

LOCUS

DEFINITION OSJNEC12117.f OSJNEC Oryza sativa (japonica cultivar-group) cDNA

clone OSJNEC12117 5', mRNA sequence.

CB657302

VERSION 1 (bases 1 to 833)

KEYWORDS

SOURCE Oryza sativa (japonica cultivar-group)

ORGANISM Oryza sativa (japonica cultivar-group)

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

Ehrhartoideae; Oryzeae; Oryza.

REFERENCE 1 (bases 1 to 833)

AUTHORS Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E.,

Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.

TITLE Large-scale identification of ESTs involved in the interaction

between rice and Magnaporthe grisea

JOURNAL Unpublished (2003)

COMMENT

Contact: Rod Wing
 Arizona Genomics Institute
 University of Arizona
 Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
 85721-0088, USA
 Tel: 520 626 3967
 Fax: 520 621 9288
 Email: http://genome.arizona.edu
 PCR Primers

FORWARD: gta aaa cga cgg cca gtg
 BACKWARD: gga aac agc tat gac cat g

Plate: 12 row: 1 column: 17

Seq primer: gta aaa cga cgg cca gtg.

FEATURES

Location/Qualifiers

1..833
 /organism="Oryza sativa (japonica cultivar-group)"
 /mol_type="mRNA"
 /cultivar="Nipponbare"
 /db_xref="taxon:39947"
 /clone="OSJNEC12117"
 /tissue_type="Leaf"
 /dev_stage="3 week"
 /lab_host="DH10B"
 /clone_lib="OSJNEC"
 /note="Vector: pBluescript II KS +; Site 1: EcoRI; Site 2:
 XhoI; 6 hrs after inoculation with Rice Blast (C9240-1)"

ORIGIN

Alignment Scores:
 Pred. No.: 1.65e-103 Length: 833
 Score: 1091.50 Matches: 227
 Percent Similarity: 91.12% Conservative: 9
 Best Local Similarity: 87.64% Mismatches: 14
 Query Match: 71.53% Indels: 9
 DB: 6 Gaps: 4
 US-10-660-226-10 (1-305) x CB657302 (1-833)

Qy 1 MetGluAlaGlyGlyValGlyLeuAlaLeuGlnAlaArgAlaAlaGlyPhe---GlySer 19
 Db 65 ATGGAGGCG---GGCGTGGGCTGGCGCTGCAGTCGCGGGCGGGGGTTCGGCGGCTCC 121
 Qy 20 SerArgHisArgGlyGlyLeu-----GlnAlaProThrGlySerLeuArg 34
 Db 122 GACCGCGCGGAGCGGCTCTACGGCGGAGGGCGGCGCGGATCGGAGCTTGAGG 181
 Qy 35 ValAlaAspProAlaGlyProAlaValAlaValAlaArgAlaArgGlySerLysProValAla 54
 Db 182 GTGCTGAGCGCGGCTGGCGGAGGCCGCTGTGTGGGCTCGCGGGTCCAAGCGGTCGCC 241
 Qy 55 ProLeuArgLeuArgAlaLysLysSerSerGlyGlyHisGluAsnSerHisAsnSerVal 74
 Db 242 CCG-----CTCCGTCGCAAGAAATCGTCGGAGGTCATGAACATTCATTAACCTCGTT 295
 Qy 75 AspGluAlaLeuLeuLeuLysArgLysSerGluGluValLeuPheThrLeuAsnGlyArg 94
 Db 296 GATGAAGCCCTCTGCTAAAGAGAAATCAGAAGAAGTTCTCTTCTATTGTAATGGACGG 355
 Qy 95 CysIleThrLeuValGlyMetMetGlySerGlyLysSerThrValGlyLysIleMetSer 114
 Db 356 TGTATTTACCTAGTTGGAATGATGGGTTCTGGAAAAAGTACTGTGGAAAAAGATCATGTCT 415
 Qy 115 GluValLeuGlyThrSerPhePheAspSerAspLysLeuValGluGlnAlaValGlyMet 134
 Db 416 GAAGTTTGGGTTATTCGTTCTTTGATAGTAAATTTGGTGAACAGCTGTGGGCATG 475
 Qy 135 ProSerValAlaGlnIlePheLysValHisSerGluAlaPhePheArgAspAsnGluSer 154
 Db 476 CCTTCAGTCGCTCAAAATTTCAAGGTTTCATAGTGAAGCTCTCTTTAGGATAATGAGAGT 535
 Qy 155 SerValLeuArgAspLeuSerSerMetArgLeuValValAlaThrGlyGlyAla 174
 Db 536 AGTGTCTTGAGGGATTTCCTCCTCAATGAAGCGATTAGTTGTTGCTACTGGAGGTGCTCT 595

Tel: 520 626 3967
Fax: 520 621 9288
Email: http://genome.arizona.edu

PCR Primers

FORWARD: gta aac cga cgg cca gta
BACKWARD: gga aac agc tat gac cat g
Plate: 11 row: D column: 04
Seq primer: gta aac cga cgg cca gta.

FEATURES

source
1..809
Location/Qualifiers
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="OSUNEC11D04"
/tissue_type="Leaf"
/dev_stage="3 week"
/lab_host="DH105"
/clone_lib="OSUNEC"
/notes="Vector: pBluescript II KS +; Site_1: EcoRI; Site_2: XhoI; 6 hrs after inoculation with Rice Blast (C9240-1)"

ORIGIN

Alignment Scores:
Pred. No.: 1.66e-98 Length: 809
Score: 1043.50 Matches: 217
Percent Similarity: 90.00% Conservative: 8
Best Local Similarity: 86.80% Mismatches: 16
Query Match: 68.38% Indels: 9
DB: 6 Gaps: 4

US-10-660-226-10 (1-305) x CB656619 (1-809)

QY 1 MetGluAlaGlyValGlyLeuAlaLeuGlnAlaArgAlaGlyPhe---GlySer 19
DB 68 ATGGAGCG---GGCGTGGGCTGGCTGCAGTCGGGGCGGGGTTTCGGCGCTCC 124
QY 20 SerArgHisArgGlyGlyLeu-----GlnAlaProThrGlySerLeuArg 34
DB 125 GACCGCGCGGAGCGGCTCTACCGCGCGGCGGGCGGGGATCGGAGCTTGAGG 184
QY 35 ValAlaAspProAlaGlyProAlaValAlaValArgAlaArgGlySerLysProValAla 54
DB 185 GTGCGCTGAGCGCGCGGTGGCGAGCGCGCTGTGTGGGCTCGCGGGTCCAAGCGGTCGCC 244
QY 55 ProLeuArgLeuArgAlaLysLysSerSerGlyGlyHisGluAsnSerHisAsnSerVal 74
DB 245 CCG-----CTCCGTCCCAAGAAATCGTCGGAGGTCATGAACATTGCACTACCTCGTT 298
QY 75 AspGluAlaLeuLeuLysArgLysSerGluGluValLeuPheTyrLeuAsnGlyArg 94
DB 299 GATGAAGCCCTCTTCTAAGAGAAATCAGAAGAAAGTTCTCTTCTATTATTTGAATGGACGG 358
QY 95 CysIleTyrLeuValGlyMetMetGlySerGlySerThrValGlyLysIleMetSer 114
DB 359 TGTATTACCTAGTTGGAATGATGGGTTCTGGAAGAAAGTACTGTGGGAAAGATCATGTCT 418
QY 115 GluValLeuGlyTyrSerPhePheAspSerAspLysLeuValGluGlnAlaValGlyMet 134
DB 419 GAAGTTTGGTATTTCGTTCTTTAGTAGTAAATTTGTTCAACACAGCTGTGGGCATG 478
QY 135 ProSerValAlaGlnIlePheLysValHisSerGluAlaPhePheArgAspAsnGluSer 154
DB 479 CCTTCAGTCGCTCAAAATTTTCAAGGTTTCATAGTGAAGCCTTCTTTAGGCAATAATGAGAGT 538
QY 155 SerValLeuArgAspLeuSerSerMetArgArgLeuValValAlaThrGlyGlyAla 174
DB 539 AGTGCTTGGGGATTGTCTCAATGAAGCGATTAGTTGTTGTCTACTCGAGGTGGTGCT 598
QY 175 ValIleArgProIleAsnTrpArgTyrMetLysArgGlyLeuSerValTrpLeuAspVal 194
DB 599 GTTATCCGACGATTAACTGGAAATACATGAAGAGGGCCTATCTGTTTGGTTGGATG 658

QY 195 ProLeuAspAlaLeuAlaArgArgIleAlaLysValGlyThrAlaSerArgProLeuLeu 214
DB 659 CCCTTGGAGCGCTCTTCTAGCGGTATTGCTAAAGTGGGACTGCTCCGCTCTCTTCTTA 718
QY 215 AspGlnProSerGlyAspProTyrAlaMetAlaPheSerLysLeuSerMetLeuAlaGln 234
DB 719 GATCAACCATCTGGTGATCCATACACAATGGCTTTTCTAAACTCAGCATGCTCGCGGAG 778
QY 235 GlnArgGlyAspAlaTyrAlaAsnAlaAsp 244
DB 779 CAAGGGCGGATGCTTTATGCAATGCTGAT 808

RESULT 6

BJ248029 689 bp mRNA linear EST 05-APR-2002
LOCUS BJ248029 Y. Ogiwara unpublished cDNA library, Wh_f Triticum
DEFINITION aestivum cDNA clone whf4j22 5', mRNA sequence.

ACCESSION BJ248029

VERSION BJ248029.1

KEYWORDS GI:20059002

SOURCE Triticum aestivum (bread wheat)

ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poideae; Triticeae; Triticum.

REFERENCE

1 (bases 1 to 689)

AUTHORS Ogiwara, Y. and Murai, K.

TITLE Expressed genes in Triticum aestivum

JOURNAL Unpublished (2002)

COMMENT

Contact: Tadao Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tsinin@genes.nig.ac.jp.

FEATURES

source

1..689
/organism="Triticum aestivum"
/mol_type="mRNA"
/cultivar="Chinese Spring"
/db_xref="taxon:4565"
/clone="whf4j22"
/tissue_type="spike at flowering date"
/dev_stage="Feekes' scale 10.5.1"
/clone_lib="Y. Ogiwara unpublished cDNA library, Wh_f"

ORIGIN

Alignment Scores:
Pred. No.: 1.29e-96 Length: 689
Score: 1024.50 Matches: 207
Percent Similarity: 93.94% Conservative: 10
Best Local Similarity: 89.61% Mismatches: 11
Query Match: 67.14% Indels: 3
DB: 4 Gaps: 2

US-10-660-226-10 (1-305) x BJ248029 (1-689)

QY 42 AlaValAlaValArgAlaArgGlySerLysProValAlaProLeuArgLeuArgAlaLys 61
DB 1 GCGGCGCGCGTGGCGGCGCGTGGCAAGCCCGCTGTCGCCG-----CTCGCGCCCAAG 54
QY 62 LysSerSer---GlyGlyHisGluAsnSerHisAsnSerValAspGluAlaLeuLeuLeu 80
DB 55 AAATCGTCCGGGGAGGTCATGAGAACTTGCACTACTCGGTGAGATGCCCTCTTGTG 114
QY 81 LysArgLysSerGluGluValLeuPheTyrLeuAsnGlyArgCysIleTyrLeuValGly 100
DB 115 AAGAAAAATCAGAAAGAGGTTCTTTCCAGTTGAATGTCGTCGTCGTCATCTACCTAGTTGA 174
QY 101 MetMetGlySerGlyLysSerThrValGlyLysIleMetSerGluValLeuGlyTyrSer 120
DB 175 ATGATGGGTTCCGGGAAAAAGCACGTCGCAAGATCTTGGCTGAAGTTTTTGGGTTATTCA 234

```

QY 121 PhePheAspSerAspLysLeuValGluGlnAlaValGlyMetProSerValAlaGlnIle 140
Db 235 TTCTTTGACAGTGATAAAATGGTGAACAAGCTGTGGCATGCCCTTCAGTTGCTCAAAAT 294

QY 141 PheLysValHisSerGluAlaPhePheArgAspAsnGluSerValLeuArgAspLeu 160
Db 295 TTCAAGTTTCACAGTGAAGCCCTTTCAGGGATATAGAGAGTAGTGTCTTGAGGAGATTG 354

QY 161 SerSerMetArgArgLeuValValAlaThrGlyGlyAlaValAlaIleArgProIleAsn 180
Db 355 TCTCAATCGCGCGATTAGTTGTTGCTACTGAGGTTGGTGTGTTATCCGACCACTTAAC 414

QY 181 TrpArgTyrMetLysArgGlyLeuSerValTrpLeuAspValProLeuAspAlaLeuAla 200
Db 415 TCGAAAAATATGAAGAAGGACTATCTGTGTTGGTGTGGATGTGCCCTTGAGGCTTTGCA 474

QY 201 ArgArgIleAlaLysValGlyThrAlaSerArgProLeuLeuAspGlnProSerGlyAsp 220
Db 475 AGCGGTATTGCTAAGTGGGACTGCTCGCTGCTCTCTCTAGATCAACCATCCGCGCAT 534

QY 221 ProTyrAlaMetAlaPheSerLysLeuSerMetLeuAlaGlnGlnArgGlyAspAlaTyr 240
Db 535 CCATACACAATGGCTTTCGAAACTCAGCACGCTCGCGAGCAAAAGGGCGACGCTTAT 594

QY 241 AlaAsnAlaAspValArgValSerLeuGluIleAlaCysLysGlnGlyHisAspAsp 260
Db 595 GCAAAATGCTGATGTGAGAGTTTCTCTCGAAGAGATCGCATCGAAGCTGGGCCATCGCAC 654

QY 261 ValSerLysLeuThrProThrAspIleAlaIle 271
Db 655 GTCCTAAGCTGACGCCGATTGATATTGCTCTG 687

RESULT 7
LOCUS CN147205
DEFINITION WOUND1_48_G03_b1_A002 Wounded leaves Sorghum bicolor cDNA clone
ACCESSION CN147205
VERSION 1
KEYWORDS 804 bp mRNA linear EST 01-APR-2004
SOURCE WOUND1_48_G03_3', mRNA sequence.
ORGANISM Sorghum bicolor (sorghum)

REFERENCE
AUTHORS Cordonnier-Pratt M.-M., Suzuki, Y., Sugano, S., Klein, R.R., Liang, C.,
Sun, F., Sullivan, R., Brady, J., Eastman, A., Miller, V., Gonzalez, M.,
Anfuso, C., Chhabra, D., Johnson, H., Kamran, D. and Pratt, L.H.
A Sorghum EST database: mechanically damaged and methyl
jasmonate-treated leaves
Unpublished (2003)
Other ESTs: WOUND1_48_G03.g1_A002
Contact: Cordonnier-Pratt MM
Laboratory for Genomics and Bioinformatics
The University of Georgia, Department of Plant Biology
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
Tel: 706 542 1860
Fax: 706 583 0210
Email: mmpratt@uga.edu
Library constructed by Dr. Yutaka Suzuki and Dr. Sumio Sugano in
the Human Genome Center, University of Tokyo Institute of Medical
Science; plant material and RNA prepared at Texas A & M University;
sequencing done in the Laboratory for Genomics and Bioinformatics,
University of Georgia. Sequence ends have been trimmed to exclude
vector and regions below Phred quality 16. Three-prime sequences
are presented as their reverse complement and have been trimmed to
exclude polyA.
Seq primer: Sug3-14 (TAGTCTAGCGCGCCGCGAC)
POLYA=Yes.
Location/Qualifiers
1. .804
FEATURES
SOURCE

```

```

/organism="Sorghum bicolor"
/mol_type="mRNA"
/cultivar="BTx623"
/db_xref="taxon:4558"
/clones="WOUND1_48_G03_A002"
/lab_host="DH10B-T1 phage-resistant E. coli"
/clone_lib="Wounded leaves"
/note="Organ: Leaf; Vector: pME18S-FL3; Site_1: XhoI;
Site_2: XhoI; The library was prepared from polyA+ RNA
harvested from 8-day-old hydroponically grown, BTx623
sorghum seedlings. For some plants, one-half of the second
leaf was crushed without damaging the midvein. For others,
methyl jasmonate was added to the growth medium to a final
concentration of 100 uM. Leaves were harvested 3 and 27 hr
after treatment and pooled. Double-stranded cDNA was
cloned unidirectionally into different DraIII sites of the
pME18S-FL3 vector (5-prime DraIII site is CACTGTGTG,
3-prime DraIII site is CACCATGTG). XhoI excises the cDNA
insert."

```

ORIGIN

```

Alignment Scores:
Pred. No.: 2,55e-95 Length: 804
Score: 1013.00 Matches: 201
Percent Similarity: 96.74% Conservative: 7
Best Local Similarity: 93.49% Mismatches: 7
Query Match: 66.38% Indels: 0
DB: 7 Gaps: 0

US-10-660-226-10 (1-305) x CN147205 (1-804)

QY 91 LeuAsnGlyArgCysIleTyrLeuValGlyMetMetGlySerGlyLysSerThrValGly 110
Db 3 TTGAACGGGAGGTGTATTACTTAGTAGGAATGATGGGTCTTGGAAAAGATACAGTGGGG 62

QY 111 LysIleMetSerGluValLeuGlyTyrSerPhePheAspSerAspLysLeuValGluGln 130
Db 63 AAGATTATGTCGAAGTCTTGGGTATTTCGTTCTTGACAGTGACAAATTAGTGAGCAA 122

QY 131 AlaValGlyMetProSerValAlaGlnIlePheLysValHisSerGluAlaPhePheArg 150
Db 123 GCTGTGGAATGCCCTTCAGTTGCTCAAAATATTCAGGTTCATAGCGAAGCCTCTTTCGG 182

QY 151 AspGluSerSerValLeuArgAspLeuSerSerMetArgArgLeuValAlaAlaThr 170
Db 183 GATAATGAGAGTAGTGTCTTGAGAGATTGTCTCCATGCAACGATTAGTTGTTGCCACC 242

QY 171 GlyGlyGlyAlaValIleArgProIleAsnTrpArgTyrMetLysArgGlyLeuSerVal 190
Db 243 GCGGTGTGTGCTGTTATCCGACCAAGTTAACTGGAATATATGAAGAGGCGCTATCTGT 302

QY 191 TrpLeuAspValProLeuAspAlaLeuAlaArgArgIleAlaLysValGlyThrAlaSer 210
Db 303 TGGTTAGGTGTGCCCTTGGATGCTCTTGCTAGGCGTATTGCTAAAGTGGGAACGCTCT 362

QY 211 ArgProLeuLeuAspGlnProSerGlyAspProTyrAlaMetAlaPheSerLysLeuSer 230
Db 363 CGTCTCTCTCCGGAACCAACCATCTGGTGATCCATACAAATGGCCCTTCTTAAGCTCAGC 422

QY 231 MetLeuAlaGlnArgGlyAspAlaTyrAlaAsnAlaAspValArgValSerLeuGlu 250
Db 423 ATGCTTGAGAGCAAAAGGGTGACGCTTATGCAATGCAGATGTAAGGGTTTCTCTAGAA 482

QY 251 GluIleAlaCysLysGlnGlyHisAspAspValSerLysLeuThrProThrAspIleAla 270
Db 483 GAGATTGCACTTAAGCAAGGTCCACATGATGTCTCTAAGCTGACACCCACCATGATCGCA 542

QY 271 IleGluSerLeuHisLysIleGluSerPheValIleGluHisThrAlaAspSerSerAla 290
Db 543 ATTGAGTCACCTTCAAGATCGAGAGCTTCGTACCGAGACACACTCTCTGATTAATCAGCT 602

QY 291 SerAspAlaGlnAlaGluSerGlnIleGlnArgIleGlnThrLeu 305

```

```

Db      603 AGCGACTCGCAAGCTGAGTCGCGAGATCCAAAGGATACAGACCTTG 647
RESULT 8
LOCUS   CB641679
DEFINITION OSJNEB01H04.f OSJNEB Oryza sativa (japonica cultivar-group) cDNA
ACCESSION CB641679
VERSION   1
KEYWORDS 797 bp mRNA linear EST 08-APR-2003
SOURCE   clone OSJNEB01H04 5', mRNA sequence.
ORGANISM Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
REFERENCE 1 (bases 1 to 797)
AUTHORS Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E.,
Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.
TITLE Large-scale identification of ESTs involved in the interaction
JOURNAL between rice and Magnaporthe grisea
COMMENT Unpublished (2003)
Contact: Rod Wing
Arizona Genomics Institute
University of Arizona
Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
85721-0088, USA
Tel: 520 626 3967
Fax: 520 621 9288
Email: http://genome.arizona.edu
PCR Primers
FORWARD: gta aaa cga cgg cca gtg
BACKWARD: gga aac agc tat gac cat g
Plate: 01 row: H column: 04
Seq primer: gta aaa cga cgg cca gtcg.

FEATURES
source
1..797
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="OSJNEB01H04"
/tissue_type="leaf"
/dev_stage="3 week"
/lab_host="DH10B"
/clone_lib="OSJNEB"
/note="Vector: pBluescript II KS +; Site 1: EcoRI; Site 2:
XhoI; 24 hrs after inoculation with Rice Blast (Che
86061)"

ORIGIN
Alignment Scores:
Pred. No.: 2,84e-95 Length: 797
Score: 1012.50 Matches: 211
Percent Similarity: 90.12% Conservative: 8
Best Local Similarity: 86.83% Mismatches: 15
Query Match: 66.35% Indels: 9
DB: 6 Gaps: 4

US-10-660-226-10 (1-305) x CB641679 (1-797)

Qy      1 MetGluAlaGlyValGlyLeuAlaLeuGlnAlaArgAlaGlyPhe---GlySer 19
Db      76 ATGGAGCGC--GGCGTGGGCTGGCGCTGCAGTCGGGGCGCGGGTTTCGGCGGCTCC 132
Qy      20 SerArgHisArgGlyGlyLeu-----GlnAlaProThrGlySerLeuArg 34
Db      133 GACCGCGCGCGAGCGCGCTCTACGCGCGCGAGGGCGCGCGGATCGCGGAGCTTGAGG 192
Qy      35 ValAlaAspProAlaGlyProAlaValAlaValArgAlaArgGlySerLysProValAla 54
Db      193 GTGCTGTAGCCGCGGTGGCGAGCGCGCTGTGTGGGCTCGCGGGTCCAAAGCCGGTCGCC 252
Qy      55 ProLeuArgLeuArgAlaLysLysSerSerGlyGlyHisGluAsnSerHisAsnSerVal 74

```

FEATURES
Location/Qualifiers
source
1..818

```

Db      253 CCG-----CTCCGTGCCAAGAAATCGTCGGAGGTCATGAACAATTGCTAATCGGTT 306
Qy      75 AspGluAlaLeuLeuLeuLeuArgLysSerGluGluValLeuPheTyrLeuAsnGlyArg 94
Db      307 GATGAAGCCCTCTTCTTAAGAGAAATCAGAAAGATTTCTCTTATTTGAATGACGG 366
Qy      95 CysIleTyrLeuValGlyMetMetGlySerGlyLysSerThrValGlyLysIleMetSer 114
Db      367 TGTATTTTACCTAGTTGGAATGATGGTTCTCGAAAAAGTACTGTGGAAAGATCATGTCT 426
Qy      115 GluValLeuGlyTyrSerPhePheAspSerAspLysLeuValGluGlnAlaValGlyMet 134
Db      427 GAAGTTTGGGTTATTCGTTCTTTGATAGTATAAATTGGTCGAAACAAGCTGTGGGCATG 486
Qy      135 ProSerValAlaGlnIlePheLysValHisSerGluAlaPhePheArgAspAsnGluSer 154
Db      487 CTTTCAGTCGCTCAAAATTTTCAGGTTTCATAGTGAAGCTCTTTTAGGGTAAATGAGAT 546
Qy      155 SerValLeuArgAspLeuSerSerMetArgArgLeuValValAlaThrGlyGlyAla 174
Db      547 AGTGTCTTTGAGGGATTGTCTCAATGAAGCGATTAGTTGTTGCTACTGGAGGTGGTCT 606
Qy      175 ValIleArgProIleAsnTyrArgTyrMetLysArgGlyLeuSerValTrpLeuAspVal 194
Db      607 GTTATCCGACCAGTTAACTGGAAATACATGAAGAGGGCCTATCTGTTGGTTGGATGG 666
Qy      195 ProLeuAspAlaLeuAlaArgArgIleAlaLysValGlyThrAlaSerArgProLeuLeu 214
Db      667 CCCTTGGACGCTCTTGTAGGCGTATTGCTAAAGTGGGACATGCTCCGCTCTTCTTCA 726
Qy      215 AspGlnProSerGlyAspProTyrAlaMetAlaPheSerLysLeuSerMetLeuAlaGln 234
Db      727 GATCAACCATCTGGTGATCCATACACAAATGGCTTTTCTANACTCAGCATGCTCGCGAG 786
Qy      235 GlnArgGly 237
Db      787 CAAAGGGGC 795

RESULT 9
LOCUS   CB684026/c
DEFINITION OSJNEF13C12.r OSJNEF Oryza sativa (japonica cultivar-group) cDNA
ACCESSION CB684026
VERSION   1
KEYWORDS CB684026.1 GI:29687751
SOURCE   EST.
ORGANISM Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
REFERENCE 1 (bases 1 to 818)
AUTHORS Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E.,
Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.
TITLE Large-scale identification of ESTs involved in the interaction
JOURNAL between rice and Magnaporthe grisea
COMMENT Unpublished (2003)
Contact: Rod Wing
Arizona Genomics Institute
University of Arizona
Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
85721-0088, USA
Tel: 520 626 3967
Fax: 520 621 9288
Email: http://genome.arizona.edu
PCR Primers
FORWARD: gta aaa cga cgg cca gtcg
BACKWARD: gga aac agc tat gac cat g
Plate: 13 row: C column: 12
Seq primer: gga aac agc tat gac cat g.

FEATURES
Location/Qualifiers
source
1..818

```

```

/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="OSJNEF13C12"
/tissue_type="Leaf"
/dev_stage="3 week"
/lab_host="DH10B"
/clone_lib="OSJNEF"
/notes="Vector: pBluescript II KS +; Site_1: EcoRI; Site_2:
XhoI; Uninfected Control"

```

ORIGIN

```

Alignment Scores:
Pred. No.:      2,28e-94      Length:      818
Score:          1004.00      Matches:    196
Percent Similarity: 96.26%      Conservative: 10
Best Local Similarity: 91.59%      Mismatches: 8
Query Match:      65.79%      Indels:    0
DB:              6          Gaps:      0

```

US-10-660-226-10 (1-305) x CB684026 (1-818)

```

QY 92 AsnGlyArgCysIleTyrLeuValGlyMetMetGlySerGlyLysSerThrValGlyLys 111
DB 818 AATGACGGTGATTACCTAGTTGGAATGATGGGTTCTGGAAGAAGTACTGTGGGAAG 759
QY 112 IleMetSerGluValLeuGlyTyrSerPhePheAspSerAspLysLeuValGluAla 131
DB 758 ATCATGCTCAAGATTGTGGGTTATTGCTTTGTAGTAGTATAAATTTGGTCGAACAAGCT 699
QY 132 ValGlyMetProSerValAlaGlnIlePheLysValHisSerClnAlaPhePheArgAsp 151
DB 698 GTGGGATGCTTCAAGTCTCAAAATTTTCAAGGTTCTAGTGAAGCCCTTCTTTAGGGAT 639
QY 152 AsnGlySerSerValLeuArgAspLeuSerSerMetArgArgLeuValValAlaThrGly 171
DB 638 AATGAGATGATGCTCTTGAGGATTTGCTCAATGAAGCGATTAGTTGCTACTGGA 579
QY 172 GlyAlaValIleArgProIleAsnTrpArgTyrMetLysArgGlyLeuSerValTrp 191
DB 578 GGTGGTGCTGTTATCCGACCACTTAACGAAATACATGAAGAAGGGCTATCTGTTGG 519
QY 192 LeuAspValProLeuAspAlaLeuAlaArgGlyIleAlaLysValGlyThrAlaSerArg 211
DB 518 TTGGATGTCCTTGGACCTCTTGCTAGCGTATTGCTAAAGTGGGGACTGCTCCCTCCGT 459
QY 212 ProLeuLeuAspGlnProSerGlyAspProTyrAlaMetAlaPheSerLysLeuSerMet 231
DB 458 CCTCTTTTAGATCAACCATCTGGTGATCCATACAAATGGCTTTTCTTAAACTCAGCATG 399
QY 232 LeuAlaGlnGlnArgGlyAspAlaTyrAlaAsnAlaAspValArgValSerLeuGlu 251
DB 398 CTCGCGAGCAAGGGCGCATGCTTATGCAAAATGCTGATGTGAGGGTTTCTCTTGAAGAG 339
QY 252 IleAlaCysLysGlnGlyHisAspValSerLysLeuThrProThrAspIleAlaIle 271
DB 338 ATTGCATCTAAACAGGGTCATGATGATGATGATGAAGCTAACCACTACCGTTGACATTCGTATT 279
QY 272 GluSerLeuHisLysIleGluSerPheValIleGluHisThrAlaAspSerSerAlaSer 291
DB 278 GAGTCGTTTCATPAAGATCCAGAACCTTTGTTCATTGAACATACCGTTGACATCCGGTTGTT 219
QY 292 AspAlaGlnAlaGluSerGlnIleGlnArgGlyIleGlnThrLeu 305
DB 218 GACTCCAGGCTGACTCAGGTGCTCAGAGGATACAGACCTTG 177

```

RESULT 10

```

CB656558          772 bp  mRNA  linear  EST 09-APR-2003
LOCUS             OSJNEC11B01.f OSJNEC Oryza sativa (japonica cultivar-group) cDNA
DEFINITION        clone OSJNEC11B01 5', mRNA sequence.
ACCESSION         CB656558

```

VERSION
KEYWORDS
SOURCE
ORGANISM

CB656558.1 GI:29660283

EST.
Oryza sativa (japonica cultivar-group)
Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.

REFERENCE

1 (bases 1 to 772)
Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E.,
Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.
Large-scale identification of ESTs involved in the interaction
between rice and Magnaporthe grisea
Unpublished (2003)

JOURNAL

COMMENT

Contact: Rod Wing

Arizona Genomics Institute

University of Arizona

Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ

85721-0088, USA

Tel: 520 626 3967

Fax: 520 621 9288

Email: http://genome.arizona.edu

PCR Primers

FORWARD: gta aaa cga cgg cca gtg

BACKWARD: gga aac agc tat gac cat g

Plate: 11 row: B column: 01

Seq primer: gta aaa cga cgg cca gtg.

FEATURES

source

1..772

/organism="Oryza sativa (japonica cultivar-group)"

/mol_type="mRNA"

/cultivar="Nipponbare"

/db_xref="taxon:39947"

/clone="OSJNEC11B01"

/tissue_type="Leaf"

/dev_stage="3 week"

/lab_host="DH10B"

/clone_lib="OSJNEC"

/note="Vector: pBluescript II KS +; Site 1: EcoRI; Site 2:

XhoI; 6 hrs after inoculation with Rice Blast (C9240-1)"

ORIGIN

Alignment Scores:

```

Pred. No.:      5,44e-93      Length:      772
Score:          990.50      Matches:    207
Percent Similarity: 89.92%      Conservative: 7
Best Local Similarity: 86.97%      Mismatches: 15
Query Match:      64.91%      Indels:    9
DB:              6          Gaps:      4

```

US-10-660-226-10 (1-305) x CB656558 (1-772)

```

QY 1 MetGluAlaGlyGlyValGlyLeuAlaLeuGlnAlaArgAlaAlaGlyPhe---GlySer 19
DB 68 ATGAGGGCG---GGCGTGGGGCTGGCGTCCAGTCCGGCGGGGGTTCGGCGGCTCC 124
QY 20 SerArgHisArgGlyGlyLeu-----GlnAlaProThrGlySerLeuArg 34
DB 125 GACCCGCCCGGAGCGCGCTCTACCGCGCGGCGGGCGCGGATCGGGAGCTTGAGG 184
QY 35 ValAlaAspProAlaGlyProAlaValAlaValAlaArgAlaArgGlySerLysProValAla 54
DB 185 GTCCGTGAGCGCGCGTGGCGAAGCCCGCTGTGTGGGCTCGCGGGTCCAAGCCGGTCGCC 244
QY 55 ProLeuArgLeuArgAlaLysLysSerSerGlyGlyHisGluAsnSerHisAsnSerVal 74
DB 245 CCG-----CTCCGTGCCAAGAAATCGTCGGAGGTCATGAACATTGTCATAACTCGGTT 298
QY 75 RepGluAlaLeuLeuLysArgLysSerGluGluValLeuPheTyrLeuAsnGlyArg 94
DB 299 GATGAAGCCCTCTCTTAAGAGAAATCAGAAAGATTCTCTTCTATTGATGACCG 358
QY 95 CysIleTyrLeuValGlyMetMetGlySerGlyLysSerThrValGlyLysIleMetSer 114

```

Db 359 TGATTACTAGTTCGAATGATGGGTTCTGGAAAAAGTACTGTGGGAAAGATCATGTCT 418

Qy 115 GluValLeuGlyTyrSerPhePheAspSerAspLysLeuValGluGlnAlaValGlyMet 134

Db 419 GAAGTTTGGGTTATTCGTTCTTTGATAGTGAATAATTTGGTGAACAAGCTGTGGGCAATG 478

Qy 135 ProSerValAlaGlnIlePheLysValHisSerGluAlaPhePheAspAspGluSer 154

Db 479 CCTTCAGTCGCTCAAAATTTTCAAGGTTTCATAGTGAAGCCTTCTTTAGGGGATATGAGAGT 538

Qy 155 SerValLeuArgAspLeuSerSerMetArgArgLeuValValAlaThrGlyGlyAla 174

Db 539 AGTGCTTCGAGGATTTGCTCAATGAAGCGGATTTAGTTGTGCTACTCGAGGTGGTGTCT 598

Qy 175 ValIleArgProIleAsnTyrMetLysArgGlyLeuSerValTyrLeuAspVal 194

Db 599 GTATCCGACCACTTAATCTGGAATATACATGAAGAGGGCTATCTGTTGGTGGATGG 658

Qy 195 ProLeuAspAlaLeuAlaArgIleAlaLysValGlyThrAlaSerArgProLeuLeu 214

Db 659 CCTTCGAGGCTCTTGCTAGGCGTATTGCTAAAGTGGGACTGCCTCCCGTCTCTTCTA 718

Qy 215 AspGlnProSerGlyAspProTyrAlaMetAlaPheSerLysLeuSerMetLeu 232

Db 719 GATCAACCATCTGGTATCCATACATATGAGGCTTTTCTAAACTCATATGCTC 772

RESULT 11
CB657303/c
LOCUS
DEFINITION
OSUNEC12117.r OSUNEC Oryza sativa (japonica cultivar-group) cDNA
clone OSUNEC12117 3', mRNA sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Oryza sativa (japonica cultivar-group)
Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoidae; Oryzeae; Oryza.
1 (bases 1 to 854)
Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E.,
Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.
Large-scale identification of ESTs involved in the interaction
between rice and Magnaporthe grisea
Unpublished (2003)
Contact: Rod Wing
Arizona Genomics Institute
University of Arizona
Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
85721-0088, USA
Tel: 520 626 3967
Fax: 520 621 9288
Email: http://genome.arizona.edu

PCR Primers
FORWARD: gta aaa cga cgg cca gtc
BACKWARD: gga aac agc tat gac cat g
Plate: 12 row: I column: 17
Seq primer: gga aac agc tat gac cat g.
Location/Qualifiers
1. .854
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="OSUNEC12117"
/tissue_type="Leaf"
/dev_stage="3 week"
/lab_host="DH10B"
/clone_lib="OSUNEC"
/note="Vector: pBluescript II KS +; Site 1: EcoRI; Site 2:
XhoI; 6 hrs after inoculation with Rice Blast (C9240-1)"

FEATURES
source

Alignment Scores:
Pred. No.: 5,38e-91 Length: 854
Score: 972.00 Matches: 191
Percent Similarity: 95.71% Conservative: 10
Best Local Similarity: 90.95% Mismatches: 9
Query Match: 63.70% Indels: 0
DB: 6 Gaps: 0

US-10-660-226-10 (1-305) x CB657303 (1-854)

Qy 96 lleTyrLeuValGlyMetMetClySerGlyLysSerThrValGlyLysIleMetSerGlu 115

Db 853 ATTTTACCTAGTTGGAAATGATGGGTTCTGAAAAAGTACTGTGGGAAAGATCATGTCTGAA 794

Qy 116 ValLeuGlyTyrSerPhePheAspSerAspLysLeuValGluGlnAlaValGlyMetPro 135

Db 793 GTTTTGGGTTATTCGTTCTTTGATAGTGAATAATTTGGTGAACAAGCTGTGGGCAATGCT 734

Qy 136 SerValAlaGlnIlePheLysValHisSerGluAlaPhePheArgAspAspGluSer 155

Db 733 TCAGTCGCTCAAAATTTTCAAGGTTTCATAGTGAAGCCTTCTTTAGGGATAATGAGAGTAGT 674

Qy 156 ValLeuArgAspLeuSerSerMetArgArgLeuValValAlaThrGlyGlyAlaVal 175

Db 673 GTCTTGAAGGATTTGCTCTCAATGAAGCGATTTAGTTGTGCTACTGGAGGTGGTGTCTGT 614

Qy 176 lleArgProIleAsnTyrArgTyrMetLysArgGlyLeuSerValTyrLeuAspValPro 195

Db 613 ATCCGACCACTTAATCTGGAATATACATGAAGAGGGCTATCTGTTGGTGGATGTGCC 554

Qy 196 LeuAspAlaLeuAlaArgArgIleAlaLysValGlyThrAlaSerArgProLeuLeuAsp 215

Db 553 TTGGACGCTCTTGCTAGGCGTATTGCTAAAGTGGGACTGCCTCCGCTCTCTTTAGAT 494

Qy 216 GlnProSerGlyAspProTyrAlaMetAlaPheSerLysLeuSerMetLeuAlaGlnGln 235

Db 493 CAACCATCTGGTGATCCATACACAATGGCTTTTTTAACTCAGCATGCTCCGAGCAAA 434

Qy 236 ArgGlyAspAlaTyrAlaAsnAlaAspValArgValSerLeuGluGluIleAlaCysLys 255

Db 433 AGGGCGGATGCTTATGCAATGCTGATGTAGGGTTTCTTTGAAGAGNTGCATCTAAA 374

Qy 256 GlnGlyHisAspAspValSerLysLeuThrProThrAspIleAlaIleGluSerLeuHis 275

Db 373 CAGGCTCATGATGATGATTCGAAGCTAACACCACTGACATTCCTATTGAGTCGTTTCAT 314

Qy 276 LysIleGluSerPheValIleGluHisThrAlaAspSerSerAlaSerAspAlaGlnAla 295

Db 313 AAGATCGAAGAACTTTGTCATTGAACATACCGTTGCAATCCGGTTGGTGACTCCAGGCT 254

Qy 296 GluSerGlnIleGlnArgIleGlnThrLeu 305

Db 253 GACTCACGCTGCTCAGAGGATACAGACCTTG 224

RESULT 12
BJ472228

LOCUS
DEFINITION

BJ472228 K. Sato unpublished cDNA library, cv. Haruna NiJo adult,
heading stage top three leaves Hordeum vulgare subsp. vulgare cDNA
clone baal32015 5', mRNA sequence.
BJ472228.1 GI:21150731

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Hordeum vulgare subsp. vulgare
Hordeum vulgare subsp. vulgare
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poideae; Triticeae; Hordeum.
1 (bases 1 to 640)
Sato,K., Saisho,D. and Takeda,K.
Barley EST sequencing project in NIG and Okayama Univ
Unpublished (2002)
Contact: Tadasu Shin-i

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.

FEATURES

source

1. .640
/organism="Hordeum vulgare subsp. vulgare"
/mol_type="mRNA"
/cultivar="Haruna Nijo"
/sub_species="vulgare"
/db_xref="taxon:112509"
/clone="baal32015"
/tissue_type="top three leaves"
/dev_stage="adult, heading stage"
/clone_lib="K. Sato unpublished cDNA library, cv. Haruna Nijo adult, heading stage top three leaves"

ORIGIN

Alignment Scores:

Pred. No.: 9,49e-91 Length: 640
Score: 968.00 Matches: 191
Percent Similarity: 96.63% Conservative: 10
Best Local Similarity: 91.83% Mismatches: 7
Query Match: 63.43% Indels: 0
DB: 4 Gaps: 0

US-10-660-226-10 (1-305) x BJ472228 (1-640)

QY 76 GluAlaLeuLeuLeuLysArgLysSerGluValLeuPheThrLeuAsnGlyArgCys 95
Db 3 GATGCCCTCTTGTGAAGAGAAATCAGAGAGGTCTTTTCAGTTGAACGGCGGTGC 62
QY 96 IleTyrLeuValGlyMetMetGlySerGlyLysSerThrValGlyLysIleMetSerGlu 115
Db 63 ATCTACCTAGTTGGAATGATGGTTCGGGMAAAGCACGCTGGGAAGATCTTGGCTGAA 122
QY 116 ValLeuGlyTyrSerPhePheAspSerAspLysLeuValGluGlnAlaValGlyMetPro 135
Db 123 GTTTTGGGTATTTCGTCTTTGACAGTGATAATTGGTCGAACAGCTCTTGGCATGCT 182
QY 136 SerValAlaGlnIlePheLysValHisSerGluAlaPhePheArgAspAsnGluSerSer 155
Db 183 TCAGTTGCTCAAAATTTTCAGGTTTCACAGTGAAGCTTCTTCAGGGATATGAGAGTAGT 242
QY 156 ValLeuArgAspLeuSerSerMetArgArgLeuValValAlaThrGlyGlyAlaVal 175
Db 243 GTCTTGAGGGATTGTCTCTATCGCGCATTAGTTGTTGCTACTGGAGGTGGTGTGT 302
QY 176 IleArgProIleAsnTrpArgTyrMetLysArgGlyLeuSerValTrpLeuAspValPro 195
Db 303 ATCCGACCAAGTTAACTGGAATAATATGAAGAAGGGCTATCTGTTTGGTTGGATGTCGC 362
QY 196 LeuAspAlaLeuAlaArgArgIleAlaLysValGlyThrAlaSerArgProLeuLeuAsp 215
Db 363 TTGGAAGCTCTTGCAGAGGGATTGCTAAAGTGGGACTGCTCGGCTCTCTTAGAT 422
QY 216 GlnProSerGlyAspProTyrAlaMetAlaPheSerLysLeuSerMetLeuAlaGlnGln 235
Db 423 CAACCATCCGGCGATCCATACACATGGCTTTTCGAAACTCAGCACCCCTCGCGGAGCAA 482
QY 236 ArgGlyAspAlaThrAlaSerAlaAspValArgValSerLeuGluGluIleAlaCysLys 255
Db 483 AGGGGCGATGCTTATGCAAAATGCTGATGTCAGAGTTTCTTTGAAGAGATTGCATCTAAG 542
QY 256 GlnGlyHisAspAspValSerLysLeuThrProThrAspIleAlaIleGluSerLeuHis 275
Db 543 CTGGGTTCATGACGATGCTCTTAAGCTGACACCAATTTGATATTGCTCTTGAGTCCGCTCCAC 602
QY 276 LysIleGluSerPheValIleGlu 283
Db 603 AAGATCGAGAGCTTTTGTCTGTCGAA 626

RESULT 13

CA146167

LOCUS

DEFINITION

SCVPRT2077E07.g RT2 Saccharum officinarum cDNA clone SCVPRT2077E07

ACCESSION

CA146167

VERSION

CA146167.1 GI:35044937

KEYWORDS

EST.

SOURCE

Saccharum officinarum

ORGANISM

Saccharum officinarum

REFERENCE

1 (bases 1 to 672)

AUTHORS

Vettore,A.L., da Silva,P.R., Kemper,E.L. and Arruda,P.

TITLE

The libraries that made SUCEST

JOURNAL

Genet. Mol. Biol. 24 (1-4), 1-7 (2001)

COMMENT

Contact: Arruda P

Centro de Biologia Molecular e Engenharia Genetica

Universidade Estadual de Campinas

Caixa Postal 6010, 13083-970, Campinas SP, Brazil

Tel: 55 19 3788 1137

Fax: 55 19 3788 1089

Email: paruda@unicamp.br

Clone distribution: clone distribution information can be found

through the Brazilian Clone Collection Center (BCCC) at

http://www.bcccenter.fcav.unesp.br

Plate: 077 row: E column: 07

Seq primer: T7 Promoter Primer.

Location/Qualifiers

1. .672

/organism="Saccharum officinarum"

/mol_type="mRNA"

/db_xref="taxon:4547"

/clone="SCVPRT2077E07"

/lab_host="DH10B"

/clone_lib="RT2"

/note="Organ: Root tips(0.3cm-long) from adult plants;

Vector: pSport1; Site 1: SalI; Site 2: NotI; An

unidirectional cDNA library generated from [root

tips(0.3cm-long) from adult plants]. cDNA was prepared

from polyA+ mRNA using SuperScript Plasmid System Kit

(Invitrogen). The double-strand cDNAs were fractionated

in a sepharose CL-2B 40cm-columns and fragments sizing

between 0.8 and 1.5 Kb were directionally cloned into the

vector. Details of each source of RNA and library

construction can be obtained at

http://sucest.lad.ic.unicamp.br/public"

ORIGIN

Alignment Scores:

Pred. No.: 2.66e-90 Length: 672
Score: 964.00 Matches: 193
Percent Similarity: 95.24% Conservative: 7
Best Local Similarity: 91.90% Mismatches: 8
Query Match: 63.11% Indels: 2
DB: 6 Gaps: 1

US-10-660-226-10 (1-305) x CA146167 (1-672)

QY 1 MetGluAlaGlyGlyValGlyLeuAlaLeuGlnAlaArgAlaAlaGlyPheGlySerSer 20
Db 47 ATGGAAGCGGGGGGGGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 106
QY 21 ArgHisArgGlyGlyLeuGlnAlaProThrGlySerLeuArgValAlaAspProAlaGly 40
Db 107 CGGCGCGGGGGGGGGCTACAGTCGCCAGCGGCGAGAGTCTGCTGGCCCGTCGGA 166
QY 41 ProAlaValAlaValArgAlaArgGlySerLysProValAlaProLeuArgLeuAla 60
Db 167 GCAGCGGTTGCTGTGCGGGCTCGCGGGTCCAGCGCGTCGCACCG-----CTCCGTGCG 220

CA146167 672 bp mRNA linear EST 24-SEP-2003
SCVPRT2077E07.g RT2 Saccharum officinarum cDNA clone SCVPRT2077E07
5', mRNA sequence.

CA146167

CA146167.1 GI:35044937

EST.

Saccharum officinarum

Saccharum officinarum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD

clade; Panicoideae; Andropogoneae; Saccharum; Saccharum officinarum

complex.

1 (bases 1 to 672)

Vettore,A.L., da Silva,P.R., Kemper,E.L. and Arruda,P.

The libraries that made SUCEST

Genet. Mol. Biol. 24 (1-4), 1-7 (2001)

Contact: Arruda P

Centro de Biologia Molecular e Engenharia Genetica

Universidade Estadual de Campinas

Caixa Postal 6010, 13083-970, Campinas SP, Brazil

Tel: 55 19 3788 1137

Fax: 55 19 3788 1089

Email: paruda@unicamp.br

Clone distribution: clone distribution information can be found

through the Brazilian Clone Collection Center (BCCC) at

http://www.bcccenter.fcav.unesp.br

Plate: 077 row: E column: 07

Seq primer: T7 Promoter Primer.

Location/Qualifiers

1. .672

/organism="Saccharum officinarum"

/mol_type="mRNA"

/db_xref="taxon:4547"

/clone="SCVPRT2077E07"

/lab_host="DH10B"

/clone_lib="RT2"

/note="Organ: Root tips(0.3cm-long) from adult plants;

Vector: pSport1; Site 1: SalI; Site 2: NotI; An

unidirectional cDNA library generated from [root

tips(0.3cm-long) from adult plants]. cDNA was prepared

from polyA+ mRNA using SuperScript Plasmid System Kit

(Invitrogen). The double-strand cDNAs were fractionated

in a sepharose CL-2B 40cm-columns and fragments sizing

between 0.8 and 1.5 Kb were directionally cloned into the

vector. Details of each source of RNA and library

construction can be obtained at

http://sucest.lad.ic.unicamp.br/public"

QY 61 LysLysSerSerGlyGlyHisGluAsnSerHisAsnSerValAspGluAlaLeuLeu 80
 Db 221 AAGAAATCGTTGGTGGTCATGAAATCTTCATAACTCCGTTGACGAAGCTCTCCTATTG 280

QY 81 LysArgLysSerGluGluValLeuPheThrLeuAsnGlyArgCysIleTyrLeuValGly 100
 Db 281 AAGAGAAATCAGAAAGAGCTTTGTTCTACTTGAACGGGAGGTATTTACTTAGTAGGA 340

QY 101 MetMetGlySerGlyLysSerThrValGlyLysIleMetSerGluValLeuGlyTyrSer 120
 Db 341 ATGATGGTCTTGAAAAAGTAGCTGGGAGAGATTATGCTGAACTCTTGGGTATTGCG 400

QY 121 PhePheAspSerAspLysLeuValGluGlnAlaValGlyMetProSerValAlaGlnIle 140
 Db 401 TTCTTCGATAGTCACAAATTAGTGAGCAAGCTGTTTGGAAATGCTTCAGTTGCCCAATA 460

QY 141 PheLysValHisSerGluAlaPhePheArgAspAsnGluSerSerValLeuArgAspLeu 160
 Db 461 TTCAGGTTTCATAGCGAAGCTCTTTTCGGGATAATGAGAGTAGTGTCTTGAGAGATNG 520

QY 161 SerSerMetArgArgLeuValAlaThrGlyGlyValAlaValIleArgProIleAsn 180
 Db 521 TCCTCCATCGCAGCATTAGTGTTGCCACCGGAGGTGGTGTGTTATCCGACCAATTAC 580

QY 181 TrpArgTyrMetLysArgGlyLeuSerValTrpLeuAspValProLeuAspAlaLeuAla 200
 Db 581 TGGGAATATATGAAGAAGGACTATCTGTTGGTTAGATGTGCCCTTGGATGCTCTNGCT 640

QY 201 ArgArgIleAlaLysValGlyThrAlaSer 210
 Db 641 AAGCGTATTGCTAAAGTGGGAAGTGCCTCT 670

RESULT 14
 AW671996 609 bp mRNA linear EST 19-JUL-2000
 LOCUS LG1_353_A11_b1_A002 Light Grown 1 (LG1) Sorghum bicolor cDNA, mRNA
 DEFINITION
 ACCESSION AW671996
 VERSION AW671996.1 GI:7535901
 KEYWORDS EST.
 SOURCE Sorghum bicolor (sorghum)
 ORGANISM Sorghum bicolor
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoideae; Andropogoneae; Sorghum.
 1 (bases 1 to 609)
 Cordonnier-Pratt,M.-M., Gingle,A., Marsala,C. and Pratt,L.H.
 An EST database from Sorghum: light-grown seedlings
 Contact: Cordonnier-Pratt MM
 Unpublished (2000)
 Laboratory for Genomics and Bioinformatics
 The University of Georgia, Department of Plant Biology
 Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
 Tel: 706 542 1860
 Fax: 706 583 0210
 Email: mmpratt@uga.edu
 Sequences have been trimmed to exclude PolyA, vector and regions
 below Phred quality 16. The threshold for highest quality sequence
 is 20.
 Seq primer: JEN REV
 High quality sequence stop: 607
 POLYA=No.

FEATURES

1..609 Location/Qualifiers
 /organism="Sorghum bicolor"
 /mol_type="mRNA"
 /db_xref="taxon:4558"
 /clone_lib="Light Grown 1 (LG1)"
 /notes="Organ: 10- to 14-day-old light-grown (greenhouse)
 seedlings; Vector: Lambda Zap; Site_1: XhoI; Site_2:
 EcoRI; The library was made from poly-A RNA in the cloning
 vector lambda ZAP II. Clones to be sequenced were

ORIGIN
 Alignment Scores:
 Pred. No.: 6 09e-90 Length: 609
 Score: 960.00 Matches: 191
 Percent Similarity: 97.51% Conservatives: 5
 Best Local Similarity: 95.02% Mismatches: 3
 Query Match: 62.91% Indels: 2
 DB: 2 Gaps: 1
 US-10-660-226-10 (1-305) x AW671996 (1-609)

QY 42 AlaValAlaValArgAlaArgGlySerLysProValAlaProLeuArgLeuAlaLys 61
 Db 12 GCGGTTGCGTGGCGGCTCGCGGTCCTCAAGCCCGTCGACCG-----CTCCGTCGCAAG 65

QY 62 LysSerSerGlyGlyHisGluAsnSerHisAsnSerValAspGluAlaLeuLeuLys 81
 Db 66 AAATCGTTCCGAGGTCATGAAAACTTCGATAACTCCGTTGACGAAGCTCTCCTGTTGAAG 125

QY 82 ArgLysSerGluGluValLeuPheThrLeuAsnGlyArgCysIleTyrLeuValGlyMet 101
 Db 126 AGGAATCCGAGAAGTCTGTTCTACTTGAACGGAGGTGTATTTACTTAGTAGGAATG 185

QY 102 MetGlySerGlyLysSerThrValGlyLysIleMetSerGluValLeuGlyTyrSerPhe 121
 Db 186 ATGGGTTCTGGAAAAAGTACAGTGGGAAGATTATGTCTGAAAGTCTTGGGTTTATTCGTTTC 245

QY 122 PheAspSerAspLysLeuValGluGlnAlaValGlyMetProSerValAlaGlnIlePhe 141
 Db 246 TTTGACAGTCACAAATTAGTGAGCAAGCTGTGTGAATGCCCTTCAGTTGCTCAATATTTC 305

QY 142 LysValHisSerGluAlaPhePheArgAspAsnGluSerSerValLeuArgAspLeuSer 161
 Db 306 AAGGTTTCATAGCAAGCTCTTTTCGGGATANTAGAGTAGTGTCTTGAGAGATTGTCTCC 365

QY 162 SerMetArgArgLeuValValAlaThrGlyGlyGlyAlaValIleArgProIleAsnTrp 181
 Db 366 TCCATGCAACGATTAGTTGTTGCCACCGCGGTGCTGTCTTATCCGACCACTTAACCTGG 425

QY 182 ArgTyrMetLysArgGlyLeuSerValTrpLeuAspValProLeuAspAlaLeuAlaArg 201
 Db 426 AATATATGAAGAAGGCGCTATCTGTTGGTTAGATGTGCCCTTCGATGCTCTTCCTAGG 485

QY 202 ArgIleAlaLysValGlyThrAlaSerArgProLeuLeuAspGlnProSerGlyAspPro 221
 Db 486 CGTATTGCTAAAGTGGGAACCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 545

QY 222 TyrAlaMetAlaPheSerLysLeuSerMetLeuAlaGlnGlnArgGlyAspAlaTyrAla 241
 Db 546 TACACAATGGCCTTCTCTAAGCTCAGCATGCTTGCAGAGCAAGGGGTGACGCTTATGCA 605

QY 242 Asn 242
 Db 606 AAT 608

RESULT 15
 CB656559/c
 LOCUS CB656559
 DEFINITION OSJNEC11B01.r OSJNEC Oryza sativa (japonica cultivar-group) cDNA
 clone OSJNEC11B01 3', mRNA sequence.
 ACCESSION CB656559
 VERSION CB656559.1 GI:29660284
 KEYWORDS EST.
 SOURCE Oryza sativa (japonica cultivar-group)
 ORGANISM Oryza sativa (japonica cultivar-group)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Ehrhartoideae; Oryzaceae; Oryza.
 1 (bases 1 to 789)
 Jantacuriyarac,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E.,
 Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.

FEATURES

789 bp mRNA linear EST 09-APR-2003
 OSJNEC11B01.r OSJNEC Oryza sativa (japonica cultivar-group) cDNA
 clone OSJNEC11B01 3', mRNA sequence.
 ACCESSION CB656559
 VERSION CB656559.1 GI:29660284
 KEYWORDS EST.
 SOURCE Oryza sativa (japonica cultivar-group)
 ORGANISM Oryza sativa (japonica cultivar-group)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Ehrhartoideae; Oryzaceae; Oryza.
 1 (bases 1 to 789)
 Jantacuriyarac,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E.,
 Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.

TITLE Large-scale identification of ESTs involved in the interaction
between rice and Magnaporthe grisea
JOURNAL Unpublished (2003)
COMMENT Contact: Rod Wing
Arizona Genomics Institute
University of Arizona
Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
85721-0088, USA
Tel: 520 626 3967
Fax: 520 621 9288
Email: <http://genome.arizona.edu>
PCR Primers
FORWARD: gta aaa cga cgg cca gtc
BACKWARD: gga aac agc tat gac cat g
Plate: 11 row: B column: 01
Seq primer: gga aac agc tat gac cat g.

FEATURES

Location/Qualifiers
1..789
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="OSUNEC11B01"
/tissue_type="Leaf"
/dev_stage="3 week"
/lab_host="DH108"
/clone_lib="OSUNEC"
/note="Vector: pBluescript II KS +; Site 1: EcoRI; Site 2:
XhoI; 6 hrs after inoculation with Rice Blast (C9240-1)"

ORIGIN

Alignment Scores:
Pred. No.: 1,79e-89 Length: 789
Score: 957.00 Matches: 188
Percent Similarity: 95.65% Conservative: 10
Best Local Similarity: 90.82% Mismatches: 9
Query Match: 62.71% Indels: 0
DB: 6 Gaps: 0

US-10-660-226-10 (1-305) x CB656559 (1-789)

Qy	99	ValGlyMetGlySerGlyLysSerThrValGlyLysIleMetSerGluValLeuGly	118
Db	788	GTGGATGATGGGTCTCGAAAAAGTACTGTGGGAAGATCATGCTGAAGTTTGGGT	729
Qy	119	TyrSerPheAspSerAspLysLeuValGluGlnAlaValGlyMetProSerValAla	138
Db	728	TATTCGTTCTTTCATAGTATATAATTGTCGAACAAGCTGTGGCATGCTTCAGTCGCT	669
Qy	139	GlnIlePheLysValHisSerGluAlaPheArgAspAsnGluSerSerValLeuArg	158
Db	668	CAAAATTTCAAGGTTTCATAGTGAAGCCTCTTTTAGGGATAATGAGAGTAGTGTCTTGAGG	609
Qy	159	AspLeuSerSerMetArgArgLeuValAlaThrGlyGlyAlaValIleArgPro	178
Db	608	GATTTGCTCTCAATGAAGCGATTAGTTGTCTACTGGAGGTGGTCTGTATCCGACCA	549
Qy	179	IleAsnTrpArgTyrMetLysArgGlyLeuSerValTrpLeuAspValProLeuAspAla	198
Db	548	GTTAACTGGAAATACATGAAGAAGGCCCTATCTGTTGGTTGGATGCGCCTTGGACGCT	489
Qy	199	LeuAlaArgArgIleAlaLysValGlyThrAlaSerArgProIleLeuAspGlnProSer	218
Db	488	CTTGCTAGGCGTATTGCTAAAGTGGGACTGCGCTCCCGCTCTCTCTTAGATCAACCATCT	429
Qy	219	GlyAspProTyrAlaMetAlaPheSerLysLeuSerMetLeuAlaGlnArgGlyAsp	238
Db	428	GGTGATCCATACACAATGCGTTTCTTAACCTCAGCATGCTCGCGAGCAAGGGCGCAT	369
Qy	239	AlaTyrAlaAsnAlaAspValArgValSerLeuGluGluIleAlaCysLysGlnGlyHis	258
Db	368	GCTTATGCAAAATGCTGATGTAGGGTTTCTCTTGAAGAGATTGCATCTAAACAGGGTCAT	309

Qy	259	AspAspValSerLysLeuThrProThrAspIleAlaIleGluSerLeuHisLysIleGlu	278
Db	308	GATGATGATCGAAGCTAACCACTGACATTTGCTATTGAGTCGTTTCATAAGATCGAG	249
Qy	279	SerPheValIleGluHisThrAlaAspSerSerAlaSerAspAlaGlnAlaGluSerGln	298
Db	248	AACTTTGTCAATTGAACATACCGTTGACAATCCGGTTGGTGACTCCAGGATGACTCACGT	189
Qy	299	IleGlnArgIleGlnThrLeu	305
Db	188	GCTCAGAGGATACAGACCTTG	168

Search completed: August 25, 2005, 23:17:18
Job time : 3688 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 25, 2005, 10:33:33 ; Search time 64 Seconds
(without alignments)
2440.377 Million cell updates/sec

Title: US-10-660-226-10
Perfect score: 1526
Sequence: 1 MEAGGVGLALQARAAGFGSS.....ADSSASDAQAESQIRIQTL 305

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Uniprot_03.*

1: uniprot_prot.*

2: uniprot_trebl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	852	55.8	287	Q7X7H9	Q7X7H9 oryza sativ
2	655	42.9	303	Q8GT76	Q8GT76 arabidopsis
3	649.5	42.6	292	1 AROK ARATH	Q9S105 arabidopsis
4	642.5	42.1	300	1 AROK LYCES	Q00497 lycopersico
5	606.5	39.7	300	2 Q8GV88	Q8GV88 arabidopsis
6	440	28.8	157	2 Q6PLR3	Q6PLR3 cucumis sat
7	429	28.1	305	2 Q9SVA4	Q9SVA4 arabidopsis
8	340.5	22.3	190	2 Q7VE85	Q7VE85 prochloroco
9	338.5	22.2	177	2 Q7NH27	Q7NH27 gloeobacter
10	338.5	22.2	181	2 Q8YXG9	Q8YXG9 anabaena sp
11	331.5	21.7	280	2 Q9LW20	Q9LW20 arabidopsis
12	318.5	20.9	189	1 AROK STNY3	P72796 synechocyst
13	312.5	20.5	191	2 Q7U469	Q7U469 synechococc
14	311.5	20.4	183	1 AROK SYNEL	Q8KH7 synechococc
15	283	18.5	192	2 Q7V904	Q7V904 prochloroco
16	272.5	17.9	185	2 Q7V3G8	Q7V3G8 prochloroco
17	267.5	17.5	244	2 Q89XW7	Q89XW7 bradyrhizob
18	262	17.2	175	2 Q7ABL5	Q7ABL5 geobacter s
19	256.5	16.8	203	2 Q6NCG8	Q6NCG8 rhodospirillum rubrum
20	254.5	16.7	210	2 Q6GIN9	Q6GIN9 bartonella
21	250	16.4	210	2 Q6G1H5	Q6G1H5 bartonella
22	249.5	16.3	168	1 AROK AQUAE	Q67925 aquifex aeo
23	246	16.1	265	2 Q655K8	Q655K8 oryza sativ
24	242.5	15.9	174	2 Q67N09	Q67N09 symbiodactyl
25	239.5	15.7	171	2 Q8RAE8	Q8RAE8 thermotoga
26	237.5	15.6	492	1 AROK THEMA	Q9WY13 thermotoga
27	235	15.4	180	2 Q6F7E4	Q6F7E4 acinetobact
28	235	15.4	211	2 Q82TC0	Q82TC0 nitrosomonas
29	234.5	15.4	200	2 Q9A435	Q9A435 caulobacter
30	228.5	15.0	184	2 Q7NZU3	Q7NZU3 chromobacte
31	227	14.9	186	2 Q7MUD6	Q7MUD6 porphyromon

32	226	14.8	291	2 Q7F2E6	Q7F2E6 oryza sativ
33	225.5	14.8	185	2 Q7VRN2	Q7VRN2 candidatus
34	224.5	14.7	192	2 Q92ME6	Q92ME6 rhizobium m
35	224	14.7	173	1 AROK YERPE	Q8ZJF7 versinia pe
36	224	14.7	173	1 AROK VIBCH	Q66419 versinia ps
37	223.5	14.6	174	1 AROK VIBCH	Q9KNV1 vibrio chol
38	222	14.5	170	2 Q6CZQ8	Q6CZQ8 erwinia car
39	220	14.4	200	2 Q8FY59	Q8FY59 brucella su
40	219.5	14.4	171	1 AROK SHEON	Q8EK20 shewanella
41	217	14.2	170	1 AROK NEIMA	P63599 neisseria m
42	217	14.2	170	1 AROK NEIMB	P63500 neisseria m
43	217	14.2	172	1 AROK SALTU	P63602 salmonella
44	217	14.2	172	1 AROK SALTU	P63601 salmonella
45	217	14.2	173	1 AROK PHOLL	Q7NA55 photorhabdu

ALIGNMENTS

RESULT 1

ID	Q7X7H9	PRELIMINARY;	PRT;	287 AA.
AC	Q7X7H9: Q7XPG2;			
DT	01-OCT-2003 (TremBLrel. 25, Created)			
DT	01-OCT-2003 (TremBLrel. 25, Last sequence update)			
DT	05-JUL-2004 (TremBLrel. 27, Last annotation update)			
DE	OSJNB0079B02.2 protein (OSJNB0003B01.21 protein).			
GN	Name=OSJNB0079B02.2; Synonyms=OSJNB0003B01.21;			
OS	Oryza sativa (japonica cultivar-group).			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;			
OC	Ehrhartoideae; Oryzaceae; Oryza.			
OX	NCBI_TaxID=39947;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	Pubmed12447439; DOI=10.1038/nature01183;			
RA	Feng Q., Zhang Y., Hao P., Wang S., Fu G., Huang Y., Li Y., Zhu J.,			
RA	Liu Y., Hu X., Jia P., Zhang Y., Zhao Q., Ying K., Yu S., Tang Y.,			
RA	Weng Q., Zhang L., Lu Y., Mu Y., Zhang L.S., Yu Z., Fan D.,			
RA	Liu X., Lu T., Li C., Wu Y., Sun T., Lei H., Li T., Hu H., Guan J.,			
RA	Wu M., Zhang R., Zhou B., Chen Z., Chen L., Jin Z., Wang R., Yin H.,			
RA	Cai Z., Ren S., Lv G., Gu W., Zhu G., Tu Y., Jia J., Zhang Y.,			
RA	Chen J., Kang H., Chen X., Sun Y., Hu Q., Zhang X., Zhang W.,			
RA	Wang L., Ding C., Sheng H., Gu J., Chen S., Ni L., Zhu F., Chen W.,			
RA	Han B.,			
RA	"Sequence and analysis of rice chromosome 4."			
RT	Nature 420:316-320(2002).			
RL	EMBL; AL606455; CAE02970.2; -.			
DR	EMBL; AL606649; CAE03630.2; -.			
DR	HSSP; P24167; 1KAG.			
DR	Gramene; Q7X7H9; -.			
DR	Gramene; Q7XPG2; -.			
DR	GO; GO:0005524; F:ATP binding; IEA.			
DR	GO; GO:0004765; F:shikimate kinase activity; IEA.			
DR	GO; GO:0008652; P:amino acid biosynthesis; IEA.			
DR	InterPro; IPR000623; ShikKinase.			
DR	PRINTS; PR01100; SHIKIMINASE.			
DR	PROSITE; PS01128; SHIKIMATE KINASE; 1.			
SQ	SEQUENCE 287 AA; 31211 MW; 3F2BFA8A03F661CE CRC64;			

Query Match 55.8%; Score 852; DB 2; Length 287;
Best Local Similarity 64.1%; Pred. No. 2.4e-57;
Matches 182; Conservative 37; Mismatches 61; Indels 4; Gaps 4;

QY	1	MEAGGVGLALQARA-AGFGSSRRHGGLOAPGTSRLRVADGAPAVARAGSKPVLRLR 59	
Db	1	MDA-VGLGRAPGAWAGLNPRRSRSRTPVPV-RFAVEKFAQPLVLGSDRRSCG-AKLKVS 57	
QY	60	AKKSGGSHNSHNSVDEALLKKKSEEVLFYNGRCIYLVGMGSGKSTVGKIMSEVLGY 119	
Db	58	CSRKPGADIKTYTYSADALVLKQAEADVVPYVPLNDRICIYLVGMGSGKTTVGKILAEVLGY 117	

```
QY 120 SPFDSDKLVQAVGMPVSAQIFKVHSEAFPRDNESVLRDLSSMRLLVATGGGAVIRPI 179
Db 118 SPFDSDKLVQAVGMPVSAQIFQVHSEAFPRDNESVLRDLSSMRLLVATGGGAVIRPI 177
QY 180 NMRYMKRGLSVLVDPLDALARRIAKVGTSASRPDLDPQSGDPYAMAFSKLSMLAQOQGDA 239
Db 178 NMSYMKKSGSTIWLVDPLDALARRIAAAGTASRPDLHQESGDPYAKAYAKLTALFQWRDSD 237
QY 240 YANADVRSLEIEACKQGHDDVSKLTPDTIAIESLHKIESFVIE 283
Db 238 YANADARVSLIEIAVKQGHSHNTVTLTPSAIAIEALLKMSFLE 281
```

RESULT 2

```
Q8GT76 PRELIMINARY; PRT; 303 AA.
AC Q8GT76;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Putative shikimate kinase.
GN Name=At2g21940;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsiis.
OX NCBI_TaxID=3702;
RN [1]
SEQUENCE FROM N.A.
RP Seki M., Iida K., Satou M., Sakurai T., Akiyama K., Iehida J.,
RA Nakajima M., Enju A., Kamiya A., Narusaka M., Carninci P., Kawai J.,
RA Hayashizaki Y., Shinozaki K.;
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK118899; BAC43483.2; -.
DR HSSP; P10880; 1SHK.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004765; F:shikimate kinase activity; IEA.
DR GO; GO:0008652; P:amino acid biosynthesis; IEA.
DR InterPro; IPR000623; Shik_kinase.
DR Pfam; PF01202; SKI; 1.
DR PRINTS; PR01100; SHIKIMTKINASE.
DR PROSITE; PS01128; SHIKIMATE_KINASE; 1.
KW Kinase.
SQ SEQUENCE 303 AA; 33980 MW; B371777C9D5F93DF CRC64;
Query Match 42.9%; Score 655; DB 2; Length 303;
Best Local Similarity 62.6%; Pred. No. 3.6e-42;
Matches 127; Conservative 36; Mismatches 40; Indels 0; Gaps 0;
```

```
QY 79 LLKRKSEVFLYNGRCIYLVGMGSGKSTVGKIMSEVLGYSFDDSKLVEQAVGMPVSA 138
Db 88 ILKRKAEVVKPYLNGRSMYLVGMGSGKTVGKLSKVILGYTFDDCDTLIEQAMNGTSVA 147
QY 139 QIFKVHSEAFPRDNESVLRDLSSMRLLVATGGGAVIRPINRWYMKRGLSVLVDPLDA 198
Db 148 EIFVHGENFFRGKETDALKLSRRYQVVVSTGGGAVIRPINWKYMKHGISIWLDPLEA 207
QY 199 LARRIAKVGTSASRPDLDPQSGDPYAMAFSKLSMLAQOQGDYANADVRSLEIEACKQGH 258
Db 208 LAHRIAAGVTSRPLLDHDSGNAYSVAFKRLSAIWDERGEAYTNANARVSLNIAAKRGY 267
QY 259 DVSKLTPDTIAIESLHKIESFV 281
Db 268 KNVSDLTPTETIAIEAFEQVLSFL 290
```

RESULT 3

```
AROK_ARATH STANDARD; PRT; 292 AA.
AC Q9SJ05;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
```

```
DE Probable shikimate kinase, chloroplast precursor (EC 2.7.1.71).
GN OrderedLocustNames=At2g21940; ORFNames=F7D8.26;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsiis.
OX NCBI_TaxID=3702;
RN [1]
SEQUENCE FROM N.A.
RP STRAIN=cv. Columbia;
RC MEDLINE=20083487; PubMed=10617197; DOI=10.1038/45471;
RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblyum T.V.,
RA Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H.L.,
RA Moffat K.S., Cronin L.A., Shen M., Pai G., Van Aken S., Umayam L.,
RA Tallon L.J., Gill J.E., Adams M.D., Carrera A.J., Creasy T.H.,
RA Goodman H.M., Somerville C.R., Copenhaver G.P., Preuss D.,
RA Nierman W.C., White O., Eisen J.A., Salzberg S.L., Fraser C.M.,
RA Venter J.C.;
RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis
thaliana.";
RL Nature 402:761-768(1999).
CC -!- CATALYTIC ACTIVITY: ATP + shikimate = ADP + shikimate 3-phosphate.
CC -!- PATHWAY: Aromatic amino acids biosynthesis; shikimate pathway;
CC fifth step.
CC -!- SUBCELLULAR LOCATION: Chloroplast (By similarity).
CC -!- SIMILARITY: Belongs to the shikimate kinase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AC007019; AAD20411.1; -.
DR PIR; A84607; A84607.
DR HSSP; P10880; 1SHK.
DR InterPro; IPR000623; Shik_kinase.
DR Pfam; PF01202; SKI; 1.
DR PRINTS; PR01100; SHIKIMTKINASE.
DR PROSITE; PS01128; SHIKIMATE_KINASE; 1.
KW Aromatic amino acid biosynthesis; ATP-binding; Chloroplast; Kinase;
KW Transferase; Transit peptide.
FT TRANSIT 1 67 Chloroplast (Potential).
FT CHAIN 68 292 Probable shikimate kinase.
FT NP_BIND 109 116 ATP (By similarity).
SQ SEQUENCE 292 AA; 32701 MW; 42C9DE79D7EE54B7 CRC64;
Query Match 42.6%; Score 649.5; DB 1; Length 292;
Best Local Similarity 62.1%; Pred. No. 9e-42;
Matches 128; Conservative 36; Mismatches 41; Indels 1; Gaps 1;
QY 79 LLKRKSEVFLYNGRCIYLVGMGSGKSTVGKIMSEVLGYSFDDSKLVEQAVGMPVSA 138
Db 88 ILKRKAEVVKPYLNGRSMYLVGMGSGKTVGKLSKVILGYTFDDCDTLIEQAMNGTSVA 147
QY 139 QIFKVHSEAFPRDNESVLRDLSSMRLLVATGGGAVIRPINRWYMKRGLSVLVDPLDA 198
Db 148 EIFVHGENFFRGKETDALKLSRRYQVVVSTGGGAVIRPINWKYMKHGISIWLDPLEA 207
QY 199 LARRIAKVGTSASRPDLDPQSGDPYAMAFSKLSMLAQOQGDYANADVRSLEIEACKQGH 258
Db 208 LAHRIAAGVTSRPLLDHDSGNAYSVAFKRLSAIWDERGEAYTNANARVSLNIAAKRGY 267
QY 259 DVSKLTPDTIAIESLHKIESFVIEH 284
Db 268 KNVSDLTPTETIAIE-VSQVLSQLLHH 292
RESULT 4
AROK_LYCES
```

```

ID AROK_LYCES STANDARD; PRT; 300 AA.
AC Q00497;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Shikimate kinase, chloroplast precursor (EC 2.7.1.71).
OS Lycopersicon esculentum (tomato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC lamiales; Solanales; Solanaceae; Solanum.
ON NCBI_TaxID=4081;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. UC82B; TISSUE=Leaf;
RX MEDLINE=93272054; PubMed=1338949;
RA Schmid J., Schaller A., Leibinger U., Boll W., Amrhein N.;
RT "The in-vitro synthesized tomato shikimate kinase precursor is
RT enzymatically active and is imported and processed to the mature
RT enzyme by chloroplasts.";
RL Plant J. 2:375-383(1992).
CC -!- CATALYTIC ACTIVITY: ATP + shikimate = ADP + shikimate 3-phosphate.
CC -!- PATHWAY: Aromatic amino acids biosynthesis; shikimate pathway;
CC fifth step.
CC -!- SUBCELLULAR LOCATION: Chloroplast.
CC -!- SIMILARITY: Belongs to the shikimate kinase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; X63560; CAA45121.1; --
DR PIR; S21584; S21584.
DR HSP; P10880; ISHK.
DR InterPro; IPR000623; Shik_kinase.
DR Pfam; PF01202; SKI; 1.
DR PRINTS; PR01100; SHIKIMTKINASE.
DR PROSITE; PS01128; SHIKIMATE_KINASE; 1.
KW Aromatic amino acid biosynthesis; ATP-binding; Chloroplast; Kinase;
FT TRANSIT 1 66 Chloroplast (Potential).
FT CHAIN 67 300 Shikimate kinase.
FT NP BIND 111 118 ATP (By similarity).
SQ SEQUENCE 300 AA; 33720 MW, 9CDAC8136922CCE3 CRC64;

Query Match 42.1%; Score 642.5; DB 1; Length 300;
Best Local Similarity 60.3%; Pred. No. 3.2e-41;
Matches 126; Conservative 35; Mismatches 47; Indels 1; Gaps 1;

QY 73 SVDEALLLRKSEEVLYLNGRCIVLVGMGSGKSTVGKIMSEVLGYSPFSDSKLVEQAV 132
DB 84 SIDEIETLNKNAEEVEYLDGRCVYLVMGCGKTTVGRILAEITLGYSPFDCDRIEQAV 143

QY 133 GMPVSAQIFKVAEAFRRNNESSVLRLDSSMRRLVATGGAVIRPINRMYKRGISVWL 192
DB 144 GGITVAEIPELRGESFRDNETVHLKLSLMHRLVVTGGAVRIPINRHHMKGISVWL 203

QY 193 DVPDLALARRIAKVGTASRLDQPSGDPYAMAFSKLSMLAQORGDAYANADVRSLEEI 252
DB 204 DVPLEALAKRITTEGKSRPLLHEESGDVYDTTLKRLTTLMETRGENYANASARVLENI 263

QY 253 ACKQGHDDVSKLTPDTIAIESLHKIESFV 281
DB 264 ALKR-EKDVCHITPAETILEVLQIENFL 291

RESULT 5
Q8GY88
ID Q8GY88 PRELIMINARY; PRT; 300 AA.
AC Q8GY88;

```

```

DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Putative shikimate kinase (Atg439540).
GN Name=Atg439540/F23K16.170;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC euroside II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Seki M., Iida K., Satou M., Sakurai T., Akiyama K., Ihida J.,
RA Nakajima M., Enju A., Kamiya A., Narusaka M., Carninci P., Kawai J.,
RA Hayashizaki Y., Shinozaki K.;
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Kim C.J., Chen H., Cheuk R., Shinn P., Bowser L., Carninci P.,
RA Chan M.M., Chang C.H., Dale J.M., Hayashizaki Y., Hsuan V.W.,
RA Ishida J., Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Lam B.,
RA Lee J.M., Lin J., Miranda M., Narusaka M., Nguyen M., Onodera C.S.,
RA Palm C.J., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,
RA Tang C.C., Toriumi M., Wong C., Wu H.C., Yanada K., Yu G., Yuan S.,
RA Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK117791; BAC42436.1; --
DR EMBL; BT005291; AAO63355.1; --
DR HSP; P10880; ISHK.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004765; F:shikimate kinase activity; IEA.
DR GO; GO:0008652; P:amino acid biosynthesis; IEA.
DR InterPro; IPR000623; Shik_kinase.
DR InterPro; IPR001680; WD40-
DR PRINTS; PR01100; SHIKIMTKINASE.
DR PROSITE; PS01128; SHIKIMATE_KINASE; UNKNOWN_1.
DR PROSITE; PS00678; WD_REPEATS_1; UNKNOWN_1.
DR Kinase.
SQ SEQUENCE 300 AA; 33777 MW, 8868B7DF6FF7D492 CRC64;

Query Match 39.7%; Score 606.5; DB 2; Length 300;
Best Local Similarity 55.6%; Pred. No. 1.9e-38;
Matches 130; Conservative 37; Mismatches 58; Indels 9; Gaps 5;

QY 57 RLRA---KKSGGCHENS---HNSVD-EALLLRKSEEVLYLNGRCIVLVGMGSGKSTV 109
DB 55 RLRSVSDKNSSALLTGTSLHSPFDEEQILKKAAEVKPYLNGRSMVLVGMGSGKTV 114

QY 110 GKIMSEVLGYSPFSDSKLVEQAVGMPVSAQIFKVAEAFRRNNESSVLRLDSSM-RLVW 168
DB 115 GKIMARSLGYTFDFCDTLIEQAMKGTVAEIPFHFGEVFRKEKEALKKLSLMYHQVV 174

QY 169 ATGGGAVIRPINRMYKRGISVWLVDLALARRIAKVGTASRLD-DOPSGDPYAMAFS 227
DB 175 STGGGAVIRPINRMYKRGISVWLVDLPLEALAHRAAAGTGRPLHDDSGDTTYAALN 234

QY 228 KLSMLAQORGDAYANADVRSLEEIACKQGHDDVSKLTPDTIAIESLHKIESFV 281
DB 235 RLSTTWDARGEATYTKASARVSLNTLKLGYRSVSDLTPAETIAIEAFEQVQSYL 288

RESULT 6
Q6PLR3
ID Q6PLR3 PRELIMINARY; PRT; 157 AA.
AC Q6PLR3;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Shikimate kinase (Fragment).
OS Cucumis sativus (Cucumber).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC euroside I; Cucurbitales; Cucurbitaceae; Cucumis.

```

```
OX NCBI_TaxID=3659;
RN [1]
RP SEQUENCE FROM N.A.
RA Kim M.S., Yang K.Y., Kim Y.C., Cho B.H.;
RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY596190; AAT02351.1; -.
DR HSSP; P24167; 1KAG.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004765; F:shikimate kinase activity; IEA.
DR GO; GO:0008652; P:amino acid biosynthesis; IEA.
DR InterPro; IPR000623; Shik_kinase.
DR Pfam; PF01202; SKI; 1.
DR PRINTS; PR01100; SHIKIMTKINASE.
DR PROSITE; PS01128; SHIKIMATE_KINASE; 1.
KW Kinase.
FT NON TER
SQ SEQUENCE 157 AA; 17514 MW; 378B1863D8F9E7F2 CRC64;

Query Match 28.8%; Score 440; DB 2; Length 157;
Best Local Similarity 58.2%; Pred. No. 5.4e-26;
Matches 85; Conservative 27; Mismatches 34; Indels 0; Gaps 0;

QY 136 SVAQIFKVFHSEAFRRDNESVLRDLSSMRRLVAVTGGGAVIRPINWRYMKRGLSVWLDVP 195
Db 3 SVAEVEKVVGEDFFREREATELRKLSLMQFVISTGGGAVTRSNWYMHKIGISVWLDVP 62

QY 196 LDALARIKVGTSAPRLDQPSGDDPYAMAFSKLSMLAQOQGDAYANADRVRSLEIACK 255
Db 63 LEALVVRISAVGTNSRPLHDSNDAYSKTIVRLSTLLEERGEAYANAEVKVSCFKIAK 122

QY 256 QGHDDVSKLTPTDIAIESLHKIESFV 281
Db 123 LGTKDVSNVTPMAIAIEAEIEETFL 148

RESULT 7
Q95VA4 PRELIMINARY; PRT; 305 AA.
AC Q95VA4;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Shikimate kinase-like protein.
GN Name=ATG39540;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Bevan M., Murphy G., Ridley P., Hudson S., Bancroft I., Mewes H.W.,
RA Lemcke K., Mayer K.F.X.;
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Murphy G., Ridley P., Hudson S., Mewes H.W., Lemcke K., Mayer K.F.X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL078620; CAB44689.1; -.
DR EMBL; AL161595; CAB80617.1; -.
DR PIR; T09370; T09370.
DR HSSP; P10880; 1SHK.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004765; F:shikimate kinase activity; IEA.
DR GO; GO:0008652; P:amino acid biosynthesis; IEA.
DR InterPro; IPR000623; Shik_kinase.
DR PRINTS; PR01100; SHIKIMTKINASE.
DR PROSITE; PS01128; SHIKIMATE_KINASE; UNKNOWN_1.
```

```
DR PROSITE; PS00678; WD_REPEATS_1; UNKNOWN_1.
KW Kinase
SQ SEQUENCE 305 AA; 34744 MW; 9C08637565F59FC CRC64;

Query Match 28.1%; Score 429; DB 2; Length 305;
Best Local Similarity 53.3%; Pred. No. 8.5e-25;
Matches 88; Conservative 28; Mismatches 47; Indels 2; Gaps 2;

QY 119 YSFFSDSKLVQAVCMPSVAQIFKVFHSEAFRRDNESVLRDLSSM-RRLVATGGGAVIR 177
Db 129 YGKIACDTLIEQAMKGTSAEIFEHFGESVFREKETEALKSLMYHQVWVSTGGGAVIR 188

QY 178 PINWRYMKRGLSVWLDVPLDALARIKAVGTASRPLL-DQPSGDDPYAMAFSKLSMLAQO 236
Db 189 PINWRYMKRGLSVWLDVPLDALARIKAVGTASRPLLHDDSGDTYTAALNRLSTIN 248

QY 237 GDAYANADRVRSLEIACKQGHDDVSKLTPTDIAIESLHKIESFV 281
Db 249 GEAYTKASARVSVLENITLKGVRVSVDUTPAEIAIEAEFEQVQSYL 293

RESULT 8
Q7VE85 PRELIMINARY; PRT; 190 AA.
AC Q7VE85;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Shikimate kinase.
GN Name=arOK; OrderedLocusNames=Pro0128;
OS Prochlorococcus marinus.
OC Bacteria; Cyanobacteria; Prochlorales; Prochlorococcaceae;
OC Prochlorococcus.
OX NCBI_TaxID=1219;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SARG / CCMF 1375 / SS120;
RX MEDLINE=22810154; PubMed=12917486; DOI=10.1073/pnas.1733211100;
RA Dufresne A., Salanoubat M., Partensky F., Attiguenave F., Axmann I.M.,
RA Barbe V., Duprat S., Galperin M.Y., Koonin E.V., Le Gall F.,
RA Makarova K.S., Ostrowski M., Oztas S., Robert C., Rogozin I.B.,
RA Scanlan D.J., Tandeau de Marsac N., Weissenbach J., Wincker P.,
RA Wolf Y.I., Hess W.R.;
RT "Genome sequence of the cyanobacterium Prochlorococcus marinus SS120,
RT a nearly minimal oxphototrophic genome.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:10020-10025 (2003).
DR EMBL; AE017161; AAP99174.1; -.
DR HSSP; P24167; 1KAG.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004765; F:shikimate kinase activity; IEA.
DR GO; GO:0008652; P:amino acid biosynthesis; IEA.
DR InterPro; IPR000623; Shik_kinase.
DR PROSITE; PS01128; SHIKIMATE_KINASE; 1.
DR Complete proteome; Kinase.
SQ SEQUENCE 190 AA; 21555 MW; 827831CEB8C141AF CRC64;

Query Match 22.3%; Score 340.5; DB 2; Length 190;
Best Local Similarity 39.8%; Pred. No. 3e-18;
Matches 80; Conservative 34; Mismatches 68; Indels 19; Gaps 4;

QY 81 KRKSEVLFLYLNRCIYLVGMWGSGKSTVGKTMSEVLGYSFSDSKLVEQAVGMPFSAQI 140
Db 4 KLRPNQVIVKLGRIIYLVGMWGSGKSTTGPHLAKLLKYSFIDQDELIEK-VAKSSVSI 62

QY 141 FKVHSEAFRRDNESVLRDLSSMRRLVAVTGGGAVIRPINWRYMKRGLSVWLDVPLDALA 200
Db 63 FREGEENGFRDIETQVLKQIGORHSLVAVTGGGLVTRSENWGLVHQGIVWLDPNRELLF 122

QY 201 RRIKAVGTASRPELLDQPSGDDPYAMAFSKLSMLAQOQGDAYANADRVRSLEIACKQGHDD 260
Db 123 ARLSKDKTVVRPLLD--NKDPKDV----LDSLIKQRYLSAEADLHISIE----- 166

QY 261 VSKLTPTDIAIESLHKIESFV 281
```

14.

```
DR EMBL; AB016889; BAB01237.1; -.
DR EMBL; BT012132; AAS76227.1; -.
DR EMBL; BT012407; AAS92323.1; -.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004765; F:shikimate kinase activity; IEA.
DR GO; GO:0008652; P:amino acid biosynthesis; IEA.
DR InterPro; IPR008162; Pyrophosphatase.
DR InterPro; IPR000623; Shik kinase.
DR PRINTS; PR01100; SHIKIMTKINASE.
DR PROSITE; PS00387; PPASE; UNKNOWN_1.
KW Kinase.
SQ SEQUENCE 280 AA; 30473 MW; D6ED9087D0AA910A CRC64;

Query Match          21.7%; Score 331.5; DB 2; Length 280;
Best Local Similarity 36.8%; Pred. No. 2.4e-17;
Matches 84; Conservative 40; Mismatches 89; Indels 15; Gaps 3;

QY 53 VAPLRRAKSSGGHSHNSHNSVDEALLKRSKEEVLFLNGRCIYLVGMGSGKSTVGRI 112
DB 50 IAPSRSLADQSASTGTSVDS-DPIDVVKRKAWDIAPELKGASIFLVGINNSIKNTGKL 108
QY 113 MEVLGYSPFSDSKLVEQAVGMPVAQIPKVHSEAFPRDNESVLRDLSSMRRLVATGG 172
DB 109 LAELRYVYFSDNLITEAAGNVSAQAALKEADKAFQSETEVLKQLSSMGLRVVCAGD 168
QY 173 GAVIRPINRWYKRGSLVWLDVPLDALARRIAKVGTASRPLDQPS--GDPYAMAFSKLS 230
DB 169 GAVQSURNLALRHGISIDWVLDITAG-----DDDSFSESPFELDTUK 216
QY 231 MLAQQRGDAYANADVRSLEEEACKQGHDDVSKLTPDTIAISLHKIE 278
DB 217 ASYKRSKGYETADVVISLEKIATKLEFEDLEAVTSEDIALSILKEIE 264
```

```
RESULT 12
AROK SYN3
ID AROK SYN3 STANDARD; PRT; 189 AA.
AC P72796;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Shikimate kinase (EC 2.7.1.71) (SK).
GN Name=arok; OrderedLocusNames=s111669;
OS Synchocystis sp. (strain PCC 6803);
OC Bacteria; Cyanobacteria; Chroococcales; Synchocystis.
OX NCBI_TaxID=1148;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97061201; PubMed=8905231;
RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
RA Miyajima N., Hiroseawa M., Sugiura M., Sasamoto S., Kimura T.,
RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K., Okumura S.,
RA Shimpou S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,
RA Tabata S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synchocystis sp. strain PCC6803. II. Sequence determination of the
RT entire genome and assignment of potential protein-coding regions.";
RL DNA Res. 3:109-136(1996).
CC -!- CATALYTIC ACTIVITY: ATP + shikimate = ADP + shikimate 3-phosphate.
CC -!- PATHWAY: Aromatic amino acids biosynthesis; shikimate pathway;
CC fifth step.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -!- SIMILARITY: Belongs to the shikimate kinase family.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; D90900; BAA16811.1; -.
```

```
DR PIR; S74659; S74659.
DR HSSP; P10880; 1SHK.
DR HAMAP; MF_00109; -.
DR InterPro; IPR000623; Shik_kinase.
DR Pfam; PF01202; SKI; 1.
DR PRINTS; PR01100; SHIKIMTKINASE.
DR PROSITE; PS01128; SHIKIMATE KINASE; 1.
KW Aromatic amino acid biosynthesis; ATP-binding; Complete proteome;
KW Kinase; Transferase.
FT NP_BIND 19 26 ATP (Potential).
SQ SEQUENCE 189 AA; 20697 MW; 41727D4EC6E585D9 CRC64;

Query Match          20.9%; Score 318.5; DB 1; Length 189;
Best Local Similarity 44.4%; Pred. No. 1.5e-16;
Matches 71; Conservative 26; Mismatches 54; Indels 9; Gaps 3;

QY 91 LNGRCIYLVGMGSGKSTVGKIMSEVLGYSPFSDSKLVEQAVGMPVAQIPKVHSEAFPR 150
DB 10 LQGVNVLFLIGMGSGKSTVGPLAQLGGRFFDADVLIERVAG-KAIADIADFEDGEDTFR 68
QY 151 DNESSVLRDLSSMRRLVATGGAVIRPINRWYKRGSLVWLDVPLDALARRIAKVGTAS 210
DB 69 DLETEVGLHLAAQKRSVIATGGVVLRRQNNWSYLHHGLVITWLDGSLLELLARLQ--GDEA 126
QY 211 RPLDQPSGDPYAMAFSKLSMLAQQRGDAYANADVRSLE 250
DB 127 RPLQVENLE-----ERLGNLLAQRPYLAQADLRFPLQ 160

RESULT 13
QYU469
ID QYU469 PRELIMINARY; PRT; 191 AA.
AC Q7U469;
DT 01-OCT-2003 (TReMBLrel. 25, Created)
DT 01-OCT-2003 (TReMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE Shikimate kinase (EC 2.7.1.71).
GN Name=arok; OrderedLocusNames=SYNW2202;
OS Synchococcus sp. (strain WH8102).
OC Bacteria; Cyanobacteria; Chroococcales; Synchococcus.
OX NCBI_TaxID=84588;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22825697; PubMed=12917641; DOI=10.1038/nature01943;
RA Palenik B., Brahamsha B., Larimer F.W., Land M.L., Hauser L.,
RA Chain P., Lamerdin J.E., Regala W., Allen E.E., McCarren J.,
RA Paulsen I.T., Dufresne A., Partensky F., Webb E.A., Waterbury J.;
RT "The genome of a motile marine Synchococcus.";
RL Nature 424:1037-1042(2003).
DR EMBL; BX569695; CAE08717.1; -.
DR HSSP; P24167; 1KAG.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004765; F:shikimate kinase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0008652; P:amino acid biosynthesis; IEA.
DR InterPro; IPR000623; Shik Kinase.
DR PROSITE; PS01128; SHIKIMATE KINASE; 1.
KW Complete proteome; Kinase; Transferase.
SQ SEQUENCE 191 AA; 20601 MW; EE012DEBC113EA97 CRC64;

Query Match          20.5%; Score 312.5; DB 2; Length 191;
Best Local Similarity 39.4%; Pred. No. 4.3e-16;
Matches 76; Conservative 31; Mismatches 71; Indels 15; Gaps 4;

QY 75 DEALLKKRSBEVLFLNGRCIYLVGMGSGKSTVGKIMSEVLGYSPFSDSKLVEQAVGM 134
DB 3 DATLSLKOR-----LGRSLYLVGMGSGKSTGTGRPLAEQLGQGFVDADAVIEQAAGC 55
QY 135 PSVAQIPKVHSEAFPRDNESVLRDLSSMRRLVATGGAVIRPINRWYKRGSLVWLDV 194
DB 56 -SIPEIFRDEGAGRALESQVLNIGQRHSLVATGGVVTQQENWGLHSGIVVWLDV 114
QY 195 PLDALARRIAKVGTASRPLDQPSGDPYAMAFSKLSMLAQQRGDAYANADVRSLEETAC 254
```



```
Db 115 VPEQLMRL--RADSTVRPLQTEPPD-----AALNALLNORRPLAEADLTVINQETP 167
Qy 255 KQGHDDVSKLTPT 267
Db 168 LAVADGILQLLPS 180

RESULT 14
AROK SYNEL
ID AROK SYNEL STANDARD; PRT; 183 AA.
AC Q8DKH7;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Shikimate kinase (EC 2.7.1.71) (SK).
GN Name=arok; OrderedLocusNames=tlr0882;
OS Synechococcus elongatus (Thermosynechococcus elongatus).
OC Bacteria; Cyanobacteria; Chroococcales; Synechococcus.
OX NCBI_TaxID=32046;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BP-1;
RX MEDLINE=22225144; PubMed=12240834;
RA Nakamura Y., Kaneko T., Sato S., Ikeuchi M., Katoh H., Sasamoto S.,
RA Watanabe A., Iriguchi M., Kawashima K., Kimura T., Kishida Y.,
RA Kiyokawa C., Kohata M., Matsumoto M., Matsuno A., Nakazaki N.,
RA Shimo S., Sugimoto M., Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the thermophilic cyanobacterium
RT Thermosynechococcus elongatus BP-1.";
RT DNA Res. 9:123-130(2002).
CC -I- CATALYTIC ACTIVITY: ATP + shikimate = ADP + shikimate 3-phosphate.
CC -I- PATHWAY: Aromatic amino acids biosynthesis; shikimate pathway;
CC fifth step.
CC -I- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -I- SIMILARITY: Belongs to the shikimate kinase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AP005371; BAC08434.1; --
CC HSSP; P10880; 1SKH.
CC HAMAP; MF_00109; --; 1
CC InterPro; IPR000623; Shik_kinase.
CC Pfam; PF01202; SKI; 1.
CC PRINTS; PR01100; SHIKIMTKINASE.
CC PROSITE; PS01128; SHIKIMATE_KINASE; 1.
KW Aromatic amino acid biosynthesis; ATP-binding; Complete proteome;
KW Kinase; Transferase.
FT NP_BIND 16 23 ATP (By similarity).
SQ SEQUENCE 183 AA; 20778 MW; 98FAE2EE114CCA32 CRC64;

Query Match 20.4%; Score 311.5; DB 1; Length 183;
Best Local Similarity 43.3%; Pred. No. 4.9e-16;
Matches 71; Conservative 28; Mismatches 52; Indels 13; Gaps 5;

Qy 91 LINGRCIYLVGMGSKSTVGKIMSEVLGYSPFSDSKLVEQAVGMPSPVAQIFKVHSEAFPR 150
Db 7 LGGANIYLVGMWGAGKTTTGRLLAQLRGLYSFVDTDAVI-TAFRQRPPIRFAGEGEPAFR 65
Qy 151 DNESSVLRLDLSMRRLVATGGAVIRPINWYMKRGLSVWLVDPLDALARRIAKVGTAS 210
Db 66 ELEQQVLAQVSSYHLLVATGGGIVLNPWNWSYLLHHGIVVWLHVLPLVLCORLRQ--DRE 123
Qy 211 RPLL-DOPSGDPYAMAFSKLSMLAQQRGDAYANAD--VRVSLEE 251
Db 124 RPLLEQPLEE-----RLGELLQARQLHYAQADLELRITLED 160
```

```
RESULT 15
Q7V904
ID Q7V904 PRELIMINARY; PRT; 192 AA.
AC Q7V904;
DT 01-OCT-2003 (TReMBLrel. 25, Created)
DT 01-OCT-2003 (TReMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE Shikimate kinase (EC 2.7.1.71).
GN Name=arok; OrderedLocusNames=PMT0163;
OS Prochlorococcus marinus (strain MIT 9313).
OC Bacteria; Cyanobacteria; Prochlorales; Prochlorococcaceae;
OC Prochlorococcus.
OX NCBI_TaxID=74547;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22825698; PubMed=12917642; DOI=10.1038/nature01947;
RA Roca G., Larimer F.W., Lamerdin J.E., Malfatti S., Chain P.,
RA Ahlgren N.A., Arellano A., Coleman M., Hauser L., Hess W.R.,
RA Johnson Z.I., Land M.L., Lindell D., Post A.F., Regala W., Shah M.,
RA Shaw S.L., Steglich C., Sullivan M.B., Ting C.S., Tolonen A.,
RA Webb E.A., Zinser E.R., Chisholm S.W.;
RT "Genome divergence in two Prochlorococcus ecotypes reflects oceanic
RT niche differentiation.";
RT Nature 424:1042-1047(2003).
DR EMBL; BX572095; CAE20338.1; --
DR HSSP; P24167; 1KAG.
DR GO; GO:0005524; F.ATP binding; IEA.
DR GO; GO:0004765; F.shikimate kinase activity; IEA.
DR GO; GO:0008652; P.amino acid biosynthesis; IEA.
DR InterPro; IPR000623; Shik_kinase.
DR PROSITE; PS01128; SHIKIMATE_KINASE; 1.
KW Complete proteome.
SQ SEQUENCE 192 AA; 20863 MW; B6F16CC52B622F76 CRC64;

Query Match 18.5%; Score 283; DB 2; Length 192;
Best Local Similarity 39.0%; Pred. No. 8e-14;
Matches 71; Conservative 27; Mismatches 64; Indels 20; Gaps 6;

Qy 91 LINGRCIYLVGMGSKSTVGKIMSEVLGYSPFSDSKLVEQAVGMPSPVAQIFKVHSEAFPR 150
Db 14 LGGRLIYLVGMWASGKSTGRPLASQLSYGFVDTDAVIEQLAQGP-IPKIFSEEGAGPR 72
Qy 151 DNESSVLRLDLSMRRLVATGGAVIRPINWYMKRGLSVWLVDPLDALARRIAKVGTAS 210
Db 73 TWESQVLAIGQRHSLVATGGGIVSKPENWGLVHGIVLWLNPGRELLRL-NADSGN 131
Qy 211 RPLLPQSGDPYAMAFSKLSMLAQQRGDAYANADVRVSLEETACKQGHDDVSKLTPTDIA 270
Db 132 RPLLQ--TEDPEA-AF---DCLFAERLPLCYEADL-----HVEVGAEBPDGIA 173
Qy 271 IE 272
Db 174 LK 175
```

Search completed: August 25, 2005, 11:12:36
Job time : 67 secs

This Page Blank (uspto)

Db 69 DLETEVLGHAAQTRSVIATGGVWLRQWNSYHLHGLVWLDGSLLELLARLQ--GDEA 126
QY 211 RPLLDQSGDPYAMAFSKLSMLAQOQGDYANADVRVLSLE 250
Db 127 RPLQVENLE-----ERLGNLLAQOQPLYAQADLRFPLQ 160

RESULT 6
A70487
shikimate kinase - Aquifex aeolicus
C:Species: Aquifex aeolicus
C:Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 09-Jul-2004
R:Accession: A70487
C:Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; Ovi
V.
Nature 392, 353-358, 1998
A:Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
A:Reference number: A70300; MUID:98196666; PMID:9537320
A:Accession: A70487
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-168 <RQF>
A:Cross-references: UNIPROT:O67925; GB:AE000776; NID:g2984355; PIDN:AAC07875.1; PID:g298
A:Experimental source: strain VFS
C:Genetics:
A:Gene: aroK
C:Superfamily: shikimate kinase; shikimate kinase homology

Query Match 16.3%; Score 249.5; DB 2; Length 168;
Best Local Similarity 33.7%; Pred. No. 6.9e-13;
Matches 62; Conservative 34; Mismatches 63; Indels 25; Gaps 5;

QY 96 IYLVGMWGSKSTVGKIMSEVLGYSFFDSKLVEQAVGMPVAQIFKVHSEAFRDNESS 155
Db 3 IYLVGMWGSKSTVGSLSLSLNIPFYDVDEEVQKREGL-SIPQIFKKGEAYFKLEPF 61
QY 156 VLRLDSMRRLVATGGGAVIRPINRYMK-RGLSYVLVDPLDALARRIAKVGTASRPLL 214
Db 62 VLKDLSEKNVISTGGGLGANEALNFMKSRGTVFIDIPFVFLERCK--DSKERPLL 119
QY 215 DQSGDPYAMAFSKLSMLAQOQGDYANADVRVLSLEIACKQGHDDVSKLTPPTDIAIESL 274
Db 120 KRP-----LDEIKNLFERRKIYSKADIKVKGK-----PPEVVVKEL 158
QY 275 HKIE 278
Db 159 LSLE 162

RESULT 7
B72389
shikimate kinase/3-dehydroquinase synthase - Thermotoga maritima (strain MSB8)
C:Species: Thermotoga maritima
C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 09-Jul-2004
C:Accession: B72389
R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey
Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.;
C.M.
Nature 399, 323-329, 1999
A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome seq
A:Reference number: A72200; MUID:99287316; PMID:10360571
A:Accession: B72389
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-492 <ARN>
A:Cross-references: UNIPROT:Q9WY13; GB:AE001715; GB:AE000512; NID:g4980839; PIDN:AAD3543
A:Experimental source: strain MSB8
C:Genetics:
A:Gene: TW0348
F:116-492/Domain: 3-dehydroquinase synthase homology <DQS>

Query Match 15.6%; Score 237.5; DB 2; Length 492;


```

A;Gene: aroK
A;Map position: 27 min
C;Superfamily: shikimate kinase; shikimate kinase homology
C;Keywords: phosphotransferase
F;70-193/Domain: shikimate kinase homology <SKI>

Query Match          14.2%; Score 216; DB 1; Length 240;
Best Local Similarity 31.6%; Pred. No. 5.5e-10;
Matches 60; Conservative 37; Mismatches 67; Indels 26; Gaps 7;

Qy 94 RCYLYGMMGSGKSTVGKIMSEVLGYSPFSDKLVQAVGMPVAQIFKVHSEAFPRDNE 153
Db 72 RNIFLVGPMGAGKSTIGRQLAQQLNMEFYDSDQIEIKRTG-ADVGWVFDLEGESEGFDRDE 130
Qy 154 SSVLRDLSSNRRLVAVATGGAV-IRPINRWYMKRGLSVWLDDVPLDALARRIAKVGTASRP 212
Db 131 EKVINELTEKQGVLATGGGSVKSRETRNRLSARGVVVYLTTIEKQLARTQR--DKKRP 188
Qy 213 LL--DOPSGDPYAMAFSKLSMLAQORGDAYAN-ADRVSLSEETACKQGHDDVSKLTPTDI 269
Db 189 LLHVEVTPPE-----VLEALANERNPLYEETADVTIRTDQSAKV-----V 229
Qy 270 AIESLHKIES 279
Db 230 ANQIIHMLSES 239

RESULT 14
shikimate kinase I [imported] - Escherichia coli (strain O157:H7, substrain RIN
C;Species: Escherichia coli
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 24-Aug-2001
C;Accession: H91157
R;Hayaashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.;
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H9
A;Reference number: A99629; MUID:21156231; PMID:11258796
A;Accession: H91157
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-240 <HAY>
A;Cross-references: GB:BA000007; PIDN:BA037655.1; PID:g13363706; GSPDB:GN00154
A;Experimental source: strain O157:H7, substrain RIMD 050952
C;Genetics:
A;Gene: EcE4232
C;Superfamily: shikimate kinase; shikimate kinase homology

Query Match          14.2%; Score 216; DB 2; Length 240;
Best Local Similarity 31.6%; Pred. No. 5.5e-10;
Matches 60; Conservative 37; Mismatches 67; Indels 26; Gaps 7;

Qy 94 RCYLYGMMGSGKSTVGKIMSEVLGYSPFSDKLVQAVGMPVAQIFKVHSEAFPRDNE 153
Db 72 RNIFLVGPMGAGKSTIGRQLAQQLNMEFYDSDQIEIKRTG-ADVGWVFDLEGESEGFDRDE 130
Qy 154 SSVLRDLSSNRRLVAVATGGAV-IRPINRWYMKRGLSVWLDDVPLDALARRIAKVGTASRP 212
Db 131 EKVINELTEKQGVLATGGGSVKSRETRNRLSARGVVVYLTTIEKQLARTQR--DKKRP 188
Qy 213 LL--DOPSGDPYAMAFSKLSMLAQORGDAYAN-ADRVSLSEETACKQGHDDVSKLTPTDI 269
Db 189 LLHVEVTPPE-----VLEALANERNPLYEETADVTIRTDQSAKV-----V 229
Qy 270 AIESLHKIES 279
Db 230 ANQIIHMLSES 239

RESULT 15
F86003
shikimate kinase I [imported] - Escherichia coli (strain O157:H7, substrain EDI
C;Species: Escherichia coli

```

